UNIVERSIDADE ESTADUAL DE SANTA CRUZ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

ESTUDO DA ESTRUTURA 3D E CANDIDATOS À INIBIDORES DA PROTEÍNA DE LIGAÇÃO À ACIL-COA (ACBP) DE *MONILIOPHTHORA RORERI*

LARISSA KAREN SILVA OLIVEIRA

ILHÉUS – BAHIA – BRASIL
MARÇO DE 2022

LARISSA KAREN SILVA OLIVEIRA

ESTUDO DA ESTRUTURA 3D E CANDIDATOS À INIBIDORES DA PROTEÍNA DE LIGAÇÃO À ACIL-COA (ACBP) DE *Moniliophthora roreri*

Dissertação apresentada à Universidade Estadual de Santa Cruz, como parte das exigências para obtenção do título de Mestre em Genética e Biologia Molecular.

Área de concentração: Genética e Biologia Molecular.

ILHÉUS – BAHIA – BRASIL

MARÇO DE 2022

O48 Oliveira, Larissa Karen Silva.
Estudo da estrutura 3D e candidatos à inibidores da proteína de ligação Acil – CoA (ACBP) de Moniliophthora roreri / Larissa Karen Silva Oliveira. – Ilhéus, BA: UESC, 2022.
xx, 103f. : il.

Orientador: Bruno Silva Andrade
Dissertação (Mestrado) – Universidade Estadual de Santa Cruz. Programa de Pós-Graduação em Genética e Biologia Molecular – PPGGBM
Inclui referências.

1. Basidiomicetos. 2. Moniliophtora roreri. 3. Proteína de ligação. 4. Bioinformática. 5. Inibidores. 6. Theobroma. 7. Candidíase. I. Título.

CDD 579.59

LARISSA KAREN SILVA OLIVEIRA

ESTUDO DA ESTRUTURA 3D E CANDIDATOS À INIBIDORES DA PROTEÍNA DE LIGAÇÃO À ACIL-COA (ACBP) DE *Moniliophthora roreri*

Dissertação apresentada à Universidade Estadual de Santa Cruz, como parte das exigências para obtenção do título de Mestre em Genética e Biologia Molecular.

Área de concentração: Genética e Biologia Molecular.

APROVADA: 31/03/2022

Prof. Dr. Aristóteles Góes Neto
(UFMG)

Prof. Eduardo Almeida Costa
(UESC)

Prof. Dr. Raner Santana
(UESC)

Prof. Dr. Bruno Silva Andrade
(Orientador - UESC/UESB)

Dedico à Honra e a Glória à Jesus Cristo. Dedico aos meus pais Sueli e Adilson, meus irmãos Brendon e Thiago e aos meus avós Gildinai e Aniel.

AGRADECIMENTOS

Agradeço, primeiramente a Deus pela dádiva da vida, por restaurar minhas forças, renovar minha fé e por todos as dificuldades superadas para a realização deste trabalho.

Agradeço ao apoio da minha família, em especial aos meus pais Suely e Adilson que sempre me incentivaram me dedicar a vida acadêmica, e o orgulho que demonstram nas minhas conquistas pessoais e profissionais; aos meus irmãos Brendon e Thiago, pelo privilégio de ser a irmã mais velha, como exemplo de dedicação e persistência.

Aos colegas da turma de mestrado, doutorado do PPGGBM - UESC, grupo proteômicos, agradeço pelos trabalhos em equipe, pelo suporte nos dias difíceis, em especial à Andria, Fernanda, Carlos Henrique, Tainá, Aline, Bruno, pelo companheirismo, e pela generosidade esclarecendo minhas dúvidas, que possibilitaram meu amadurecimento pessoal e profissional na troca de experiências e conhecimentos.

Agradeço aos professores do PPGGBM por serem atenciosos, pela troca de conhecimentos científicos e experiências pessoais. E especialmente á Mara e Fabrícia pela assistência prestada em tempo hábil e esclarecimentos de dúvidas.

Aos meus colegas do LBQC, da UESB em Jequié, pela hospitalidade, parceria e das diversas trocas de aprendizagem na área de bioinformática, na utilização das ferramentas, e da disponibilidade em ajudar superar as minhas limitações.

Ao CNPq pela disponibilização da bolsa, que me possibilitou me dedicar inteiramente ao trabalho e a suprir minhas necessidades básicas.

AGRADECIMENTO ESPECIAL

Agradeço ao meu orientador o Prof. Dr. Bruno Silva Andrade, por toda generosidade em me receber como aluna, pela orientação e toda atenção prestada durante todo esse período do mestrado, seja presencialmente ou pelas reuniões no Google meet, principalmente, pela atenção, paciência e ensinamentos. Obrigada por me permitir tornar uma profissional mais completa e por acreditar na minha capacidade de cumprir com os trabalhos propostos com todo seu suporte. Seu comprometimento e esforços com a pesquisa científica se tornou uma referência pra mim.

Ao meu co-orientador Prof. Dr. Carlos Priminho Pirovani, pela humildade de me receber e acreditar no meu potencial antes mesmo de ingressar no mestrado; por me apresentar ao meu orientador, e por ser uma referência como professor dedicado e profissional exemplar,

se fazendo presente e colaborando com troca de ensinamentos científicos nas discussões do grupo Proteômicos.

ÍNDICE

EXTRATO.....xii
ABSTRACT.....xiii
LISTA DE FIGURAS..... xiv
LISTA DE TABELAS.....xvii
LISTA DE SÍMBOLOS, SIGLAS E ABREVIATURAS.....xviii
1. INTRODUÇÃO.....1
2. OBJETIVOS.....3
2.1 Objetivo Geral.....3
2.2 Objetivos específicos.....4
3. REVISÃO DE LITERATURA.....4
3.1 *Moniliophthora roreri:* descrição, nomenclatura, taxonomia e hospedeiros.....4
3.1.2 Ciclo de vida, sintomas e características da infecção.....7
3.1.3 Origem, distribuição geográfica atual e impacto econômico.....9
3.1.4 Principais estratégias de controle no manejo integrado.....12
3.2 Uma visão geral da estrutura e função de proteínas.....14
3.3 Proteína de ligação à acil-CoA (ACBP).....16
3.3.1 Inibidores da ACBP: novas drogas com atividade anti-ACBP.....20
3.4 Utilização de ferramentas de bioinformática aplicadas na predição de estruturas proteicas e busca de inibidores.....21
3.4.1 Modelagem Molecular de Proteínas: Modelagem por Homologia, *Ab initio*, *Threading* (I-Tasser) 22
3.4.2 Triagem virtual baseada na estrutura do ligante (LBVS) e Triagem virtual baseada na estrutura do receptor (RBVS).....25
3.4.3 Acoplamento Molecular (*Molecular Docking*) 27
3.4.4 Simulações de Dinâmica Molecular (DM): Aspectos gerais e aplicações.....32
4. MATERIAL E MÉTODOS.....38
4.1 Análises *in sílico* da ACBP de *Moniliophthora roreri*.....38
4.2 Modelagem molecular por homologia da ACBP de *Moniliophthora roreri*.....39
4.3 Minimização de energia e validação da ACBP de *Moniliophthora roreri*.....40
4.4 Triagem virtual baseada no receptor (RBVS).....40
4.5 Preparação do receptor e ligante para o acoplamento molecular (*docking*).....41
4.6 Processo de Acoplamento molecular (*Docking*).....42

4.7 Modelagem por homologia da ACBP de *Theobroma cacao*.....43
4.8 Simulações de Dinâmica Molecular.....43
5. RESULTADOS E DISCUSSÃO.....44
5.1 Análise da sequência da ACBP de *Moniliophthora roreri* (MrACBP).....44
5.2 Modelagem da estrutura 3D de MrACBP.....46
5.3 Modelagem da estrutura 3D da ACBP de *Theobroma cacao* (TcACBP).....54
5.4 Identificação da alça divergente em TcACBP.....57
5.5 Triagem virtual baseada no receptor e estudo de *Docking* entre MrACBP e ligantes.....59
5.6 Estudo de *Docking* entre TcACBP e ligantes.....66
5.7 Dinâmica Molecular de estabilidade entre MrACBP e Isoxicam.....72
6. CONCLUSÃO.....77
7. PERSPECTIVAS.....78
8. REFERÊNCIAS BIBLIOGRÁFICAS……..…...78

EXTRATO

OLIVEIRA, Larissa Karen Silva, M.S., Universidade Estadual de Santa Cruz, Ilhéus, março de 2022. **Estudo da estrutura 3D e candidatos à inibidores da proteína de ligação à Acil-CoA (ACBP) de *Moniliophthora roreri*.** Orientador: Bruno Silva Andrade. Co-orientador: Carlos Priminho Pirovani.

O fungo *Moniliophthora roreri* é o agente patogênico da Monilíase, doença que afeta com grande impacto destrutivo pelo fato de atacar exclusivamente os frutos, das espécies dos gêneros *Theobroma* e *Herrania*, com prejuízos no cultivo que variam cerca de 30 a 100% na produção. Acometendo ao logo da história vários países da América do sul, América central e América do norte. Atualmente, a monilíase encontra-se em fase invasiva na América tropical e recentemente foi constatado sua detecção em cacaueiros e cupuaçuzeiros na cidade de Cruzeiro do Sul, Acre. Medidas fitossanitárias com manejo integrado e estratégias de contenção e erradicação tem sido adotada por técnicos de várias instituições, entretanto, ainda faltam estratégias de controle para intervenção desta praga que sejam eficazes, acessíveis, e aplicáveis a pequenos produtores de cacau. Devido a importância econômica do cacaueiro para a região Sul da Bahia, faz-se necessário estudos envolvendo proteômica e bioinformática que sejam relevantes para o desenvolvimento do fungo *M. roreri*, e que sejam aplicados associados as demais medidas fitossanitárias. A proteína de ligação à acil-CoA (ACBP) presente no fungo *Moniliophthora roreri*, constitui uma família de proteínas que facilitam a ligação de ésteres de acil-CoA de cadeia longa em um domínio altamente conservados em todos os organismos distribuídos nos quatro reinos eucarióticos, *Animalia*, *Plantae*, *Fungi* e *Protista*, e em algumas espécies de eubactérias, além de possuírem a capacidade de se ligarem a um éster de acil-CoA de cadeia longa com alta afinidade; os ésteres de acil-CoA de cadeia longa são intermediários no metabolismo lipídico e modulam uma ampla gama de funções celulares importantes, como regulação do metabolismo energético, síntese lipídica, transdução de sinal, regulação do fluxo iônico e expressão gênica, atuam para formar pools intracelulares de acil-CoA, transportar ésteres de acil-CoA. Esse estudo teve como objetivo principal a modelagem da estrutura tridimensional bem como, a triagem de prováveis inibidores da proteína de ligação à acil-CoA (ACBP) de *M. roreri*, utilizando metodologias computacionais de bioinformática e química computacional. Para melhor compreensão do funcionamento desta proteína, tendo como ponto de partida a construção do seu modelo tridimensional, empregamos uma série de metodologias computacionais como modelagem por homologia, triagem virtual baseada no receptor, acoplamento molecular (*docking*) e simulações de dinâmica molecular. Desse modo, foi possível predizer com confiabilidade o modelo estrutural da proteína ACBP de *M. roreri* e assim, elucidou a compreensão acerca da estrutura proteica ACBP, acredita-se que as melhores moléculas candidatas selecionadas, o ácido gambogico, isoxicam, meclociclina, cloxacilina e pregnenolona, possam atuar como inibidores em potencial desta proteína em testes *in sílico*, *in vitro* e *in vivo*, sendo imprescindível a compreensão da estrutura energética e dinâmica molecular das interações da proteína-ligante, bem como seus mecanismos de ação. Portanto, este estudo propõe novas moléculas com potencial fungicida com ação frente à ACBP de *M. roreri* aptas para prosseguirem em testes *in vitro* e *in vivo* no intuito de atuarem como novas perspectivas no controle da monilíase do cacaueiro.

Palavras-chave: *M. roreri*, ACBP, triagem computacional, inibidores, *Theobroma cacao*, Monilíase.

ABSTRACT

OLIVEIRA, Larissa Karen Silva, M.S., Universidade Estadual de Santa Cruz, Ilhéus, March 2022. **Study of the 3D structure and inhibitors candidates of the Acil-CoA binding protein (ACBP) of *Moniliophthora roreri.*** Advisor: Bruno Silva Andrade. Co-advisor: Carlos Priminho Pirovani.

The fungus *Moniliophthora roreri* is the pathogenic agent of Moniliasis, a disease that affects with great destructive impact because it attacks exclusively the fruits of the *Theobroma* and *Herrania* species, with crop losses ranging from about 30 to 100% in production. Affecting throughout history several countries in South America, Central America and North America. Currently, moniliasis is in an invasive phase in tropical America and its detection was recently confirmed in cocoa and cupuaçu trees in the city of Cruzeiro do Sul, Acre. Phytosanitary measures with integrated management and containment and eradication strategies have been adopted by technicians from several institutions, however, there is still a lack of control strategies for the intervention of this pest that are effective, accessible, and applicable to small cocoa producers. Due to the economic importance of cocoa for the southern region of Bahia, it is necessary to carry out studies involving proteomics and bioinformatics that are relevant for the development of the M. roreri fungus, and that are applied in association with other phytosanitary measures. The acyl-CoA binding protein (ACBP) present in the fungus *Moniliophthora roreri*, constitutes a family of proteins that facilitate the binding of long-chain acyl-CoA esters in a highly conserved domain in all organisms distributed in the four eukaryotic kingdoms, *Animalia*, *Plantae*, *Fungi* and *Protista*, and in some species of eubacteria, in addition to having the ability to bind a long-chain acyl-CoA ester with high affinity; long-chain acyl-CoA esters are intermediates in lipid metabolism and modulate a wide range of important cellular functions such as regulation of energy metabolism, lipid synthesis, signal transduction, regulation of ion flux and gene expression, act to form intracellular pools of acyl-CoA, carry esters of acyl-CoA. This study had as main objective the modeling of the three-dimensional structure as well as the screening of probable inhibitors of the acyl-CoA binding protein (ACBP) of *M. roreri*, using computational methodologies of bioinformatics and computational chemistry. To better understand the functioning of this protein, having as a starting point the construction of its three-dimensional model, we employ a series of computational methodologies such as homology modeling, receptor-based virtual sorting, molecular docking and molecular dynamics simulations. Thus, it was possible to reliably predict the structural model of the ACBP protein of *M. roreri* and, thus, clarified the understanding of the ACBP protein structure, it is believed that the best candidate molecules selected, gambogic acid, isoxicam, meclocycline, cloxacillin and pregnenolone, can act as potential inhibitors of this protein in *in sílico*, *in vitro* and *in vivo* tests, being essential to understand the energetic structure and molecular dynamics of the protein-ligand interactions, as well as their mechanisms of action. Therefore, this study proposes new molecules with fungicidal potential with action against the ACBP of *M. roreri* able to continue *in vitro* and *in vivo* tests in order to act as new perspectives in the control of cocoa thrush.

Key-words: *M. roreri*, ACBP, computer screening, inhibitors, *Theobroma cacao*, Moniliasis.

LISTA DE FIGURAS

Figura 1: Microscopia eletrônica de varredura (SEM) e transmissão (TEM) para mostrar detalhes taxonômicos críticos de *Moniliophthora roreri*.....6
Figura 2: Infecção por monilíase *(M. roreri)* em parentes do cacau.....7
Figura 3: Ciclo da doença de *Moniliophthora roreri* em *Theobroma cacao*.....8
Figura 4: Sintomatologia de *Moniliophthora roreri* nas fases biotrófica e necrotrófica inicial....9
Figura 5: Distribuição geográfica da Monilíase do cacau causada por *Moniliophthora roreri*.....11
Figura 6: Níveis de estruturas nas proteínas.....14
Figura 7: Comparação de sequências de aminoácidos de ACBPs de espécies diferentes.....17
Figura 8: Estrutura de OsACBP1 (*Oryza sativa*) na representação de fita colorida.....18
Figura 9: Estrutura da ACBP em *Moniliophthora perniciosa*.....19
Figura 10: MES vinculado a *M. perniciosa* ACBP.....20
Figura 11: Esquema geral do método teórico da modelagem comparativa.....23
Figura 12: Esquematização das abordagens TBVS e LBVS.....27
Figura 13: Representação esquemática da formação do acoplamento molecular (*docking*) entre uma proteína-alvo interagindo com o ligante e gerando um complexo, com a região do sítio ativo destacado em vermelho.....28
Figura 14: Equações da abordagem no AutoDock.....30
Figura 15: Equação da função de pontuação do AutoDock Vina.....31
Figura 16: Construção da "caixa" (*grid box*) com coordenadas de x, y e z do sítio ativo do receptor MrACBP para realizar o acoplamento molecular (*docking*).....32
Figura 17: Equação para aproximar as forças atômicas que atuam em cada átomo e que governam o movimento molecular.....33
Figura 18: A. Esquematização da abordagem da simulação de Dinâmica Molecular. B. Algoritmo básico de DM.....34
Figura 19: Ilustração da variedade de processos biomoleculares dinâmico ao longo do tempo que podem ser observados a partir das simulações de DM.....36
Figura 20: Aplicações de Simulações de Dinâmica Molecular nos estudos estruturais e dinâmicos de uma biomolécula, avaliando sua flexibilidade conformacional e estabilidade.....37

Figura 21: Aplicações de simulações de Dinâmica Molecular com alguma perturbação ou mudança controlada no sistema biomolecular.....38
Figura 22: Representação da construção das coordenadas de x, y e z do sítio ativo do receptor (a) MrACBP e (b)TcACBP para realizar o acoplamento molecular (*docking*).....42
Figura 23: Resultado da predição do peptídeo sinal da proteína MrACBP.....45
Figura 24: Domínio e superfamília da proteína ACBP de *M. roreri*.....46
Figura 25: Estrutura tridimensional da ACBP de *Moniliophthora perniciosa*.....48
Figura 26: Resultado da busca de modelos homólogos no SWISS MODEL.....49
Figura 27: Estrutura tridimensional da ACBP de *Moniliophthora roreri* (modelo alvo) 49
Figura 28: Alinhamento estrutural entre MrACBP com energia minimizada – modelo construído (verde) e MpACBP (PDB 3FP5) - molde (lilás).....50
Figura 29: Alinhamento entre as sequências de aminoácidos da MrACBP e MpACBP (PDB 3FP5) 50
Figura 30: Esquema de avaliação e validação do modelo homólogo PDB 3FP5.....52
Figura 31: Gráfico de Ramachandran, indicando boa qualidade do modelo construído (acima de 90%) 53
Figura 32: a e b. Representação estrutural (3D) da MrACBP (verde) com região do sítio ativo destacado (vermelho) composto pelos resíduos de aminoácidos: GLN 14, TYR 33, TYR 36, LYS 37, LYS 59, TYR 78.....54
Figura 33: Estrutura tridimensional de TcACBP (*Theobroma cacao*) 55
Figura 34: Alinhamento entre as sequências de aminoácidos da TcACBP e OsACBP (PDB 5H3I) 56
Figura 35: Alinhamento estrutural entre TcACBP – energia minimizada (rosa) e o molde OsACBP PDB ID 5H3I (cinza).....56
Figura 36: Validação do modelo TcACBP.....57
Figura 37: Alinhamento múltiplo entre as sequências de aminoácidos de TcACBP, MpACBP e MrACBP.....58
Figura 38: Alinhamento estrutural entre TcACBP (rosa), MpACBP (lilás) e MrACBP (verde).....58
Figura 39: Estruturas 3D do TOP 5 inibidores selecionados com maior energia de afinidade.....61
Figura 40: Complexo formado pela proteína ACBP de *M. roreri* (verde) e o ácido gambogico (ligante CID 9852185) 62

Figura 41: Complexo formado pela proteína ACBP de *M. roreri* (verde) e o isoxicam (ligante CID 54677972) 63
Figura 42: Complexo formado pela proteína ACBP de *M. roreri* (verde) e o Meclociclina (ligante CID 54676539) 64
Figura 43: Complexo formado pela proteína ACBP de *M. roreri* (verde) e o Cloxacilina (ligante CID 6098) 65
Figura 44: Complexo formado pela proteína ACBP de *M. roreri* (verde) e o Pregnenolona (ligante CID 8955) 66
Figura 45: Complexo formado pela proteína TcACBP (rosa) e o Pregnenolona (ligante CID 8955) 68
Figura 46: Complexo formado pela proteína TcACBP (rosa) e o Cloxacilina (ligante CID 6098) 69
Figura 47: Complexo formado pela proteína TcACBP (rosa) e o Ácido gambógico (ligante CID 9852185) 70
Figura 48: Complexo formado pela proteína TcACBP (rosa) e o Isoxicam (ligante CID 54677972) 72
Figura 49: Gráfico RMSD (derivação da raiz quadrática média) da dinâmica de estabilidade entre MrACBP e o ligante Isoxicam.....73
Figura 50: Gráfico do Potencial de Energia (Coulomb) da dinâmica de estabilidade entre MrACBP e o ligante Isoxicam.....74
Figura 51: Gráfico do Potencial de Energia (Lennard-Jones) da dinâmica de estabilidade entre MrACBP e o ligante Isoxicam.....75
Figura 52: Gráfico da dinâmica molecular de estabilidade entre MrACBP e o ligante Isoxicam – Ligação de hidrogênio.....76
Figura 53: Gráfico da dinâmica molecular de estabilidade entre MrACBP e o ligante Isoxicam – Desvio da raiz quadrática média de flutuação.....77

LISTA DE TABELAS

Tabela 1 - Os 28 principais compostos para inibição da ligação da proteína de ligação à acil Coa graxo (CpACBP1) para NBD-palmitoil-CoA, e as bioatividades principais.....21
Tabela 2 - Resultados da busca por ACBP de *Moniliophthora roreri* no banco de dados de proteínas (NCBI) e pesquisa de alinhamento local por estruturas homólogas na ferramenta do BLASTP.....46
Tabela 3 - Lista de inibidores contra a proteína de ligação de acil-CoA graxo (ACBP) anti-*Cryptosporidium parvum*, compostos semelhantes, e drogas conhecidas a partir do banco de dados DrugBank.....59
Tabela 4 - Ligantes positivos que realizaram *docking* molecular com a MrACBP (receptor) no AutoDock Vina e que foram complexados com o receptor no PyMol.....60
Tabela 5 - Ligantes positivos que realizaram *docking* molecular com a TcACBP (receptor) no AutoDock Vina e que foram complexados com o receptor no PyMol.....67
Tabela 6 - Comparação da afinidade entre os complexos de MrACBP e os cinco ligantes de maior interação e suas energias de afinidade, com TcACBP e os mesmos ligantes.....70

LISTA DE SÍMBOLOS, SIGLAS E ABREVIATURAS

ACB – do inglês, *Binding Acil CoA* (ligação a acil CoA)

ACBP – do inglês, *Protein binding acil CoA* (proteína de ligação à acil-coenzima A)

CoA – Coenzima A

MES – ácido 2- (N-morfolino) etanossulfônico

α – alfa, letra do alfabeto grego

β – beta, letra do alfabeto grego

π – letra grega, que representa o número Pi

CpACBP1 – proteína de ligação de acil-CoA graxo de *Cryptosporidium parvum*

HCl – ácido clorídrico

M. roreri – Moniliophthora roreri

T. cacao – Theobroma cacao

RBVS – do inglês, *Receptor-Based Virtual Screening* (Triagem Virtual Baseada no Receptor)

TBVS – do inglês, *Target-Based Virtual Screening* (Triagem Virtual Baseada em Alvos)

LBVS – do inglês, *Ligand-Based Virtual Screening* (Triagem Virtual Baseada em Ligantes)

C. perniciosa – Crinipellis perniciosa

SEM – Microscopia eletrônica de varredura

TEM – Microscopia eletrônica de transmissão

AFLP – do inglês, *Amplified Fragment Length Polymorphism* (polimorfismo de comprimento de fragmento amplificado)

ISSR – do inglês, *Inter Simple Sequence Repeat* (Repetição de Sequência Inter Simples)

ITS – do inglês, *Internal transcribed spacer* (Região Interna Transcrita)

MAPA - Ministério da Agricultura Pecuária e Abastecimento

PNPV – Programa Nacional de Prevenção e Vigilância de Pragas Quarentenárias Ausentes

N-terminal – amino-terminal

C-terminal – carboxi-terminal

kDa – Kilodalton

DBI – do inglês, *Diazepam binding inhibitor* (inibidor de ligação do Diazepam)

EP – do inglês, *Endozepines* (Endozepina)

Å – angstrom

3D – tridimensional

RCSB – do inglês, *Research Collaboraty for Structural Bioinformatics* (Laboratório de Pesquisa em Bioinformática Estrutural)

PDB – do inglês, *Protein Data Bank* (Banco de Dados de Proteínas)

AIDS – do inglês, *Acquired Immunodeficiency Syndrome* (Síndrome da Imunodeficiência Humana)

NBD – Nitrobenzoxadiazol (marcador fluorescente)

MM – Modelagem Molecular

NMR – do inglês, *Nuclear Magnetic Resonance* (Ressonância Magnética Nuclear)

pH – Potencial de Hidrogeniônico

PDBQT – formato de entrada no AutoDock Tools

MrACBP – Proteína de ligação à acil CoA de *Moniliophthora roreri*

TcACBP – Proteína de ligação à acil CoA de *Theobroma cacao*

MpACBP- Proteína de ligação à acil CoA de *Moniliophthora perniciosa*

OsACBP - Proteína de ligação à acil CoA de *Oryza sativa*

DM – Dinâmica Molecular

MM – Mecânica Molecular

Fs – femtossegundos

HPC – do inglês, *High Performance Computing* (computação de alto desempenho)

Compute Pi/Mw Expasy – do inglês, *Compute Pi (isoelectric point) and Mw (molecular weight) Expasy* (Calcule o ponto isoelétrico e peso molecular)

CDD – do inglês, *Conserved Domains Database* (Banco de dados de domínios conservados)

NCBI – do inglês, *National Center for Biotechnology Information* (Centro Nacional de Informações sobre Biotecnologia)

BLASTP – do inglês, *Basic Local Alignment Search Tool* (Ferramenta básica de pesquisa de alinhamento local de proteínas)

T-COFFEE – do inglês, T-COFFEE *Multiple Sequence Alignment Server Program* (programa de alinhamento de múltiplas sequências)

RMSD – do inglês, *Root-Mean-Square Desviation* (Desvio da Raiz Média Quadrática ou Desvio Médio Quadrado)

ID – Código de identificação

SMILES – do inglês, *Simplified Molecular Input Line Entry Specification* (Especificação da linha de entrada molecular simplificada)

PLIP Web Tool – do inglês, *Protein-Ligand Interaction Profiler* (Perfilador de interação proteína-ligante)

FASTA – formato em códigos de uma única letra para representar as sequências de aminoácidos (proteínas)

GMQE – do inglês, *Global Model Quality Estimation* (Estimativa de Qualidade Global do Modelo)

QMEAN – do inglês, *Quality Model Energy Analysis* (Análise Energética da Qualidade do Modelo/ Estimativa de Qualidade Globais e Locais)

aa – resíduos de Aminoácidos

Q / GLN – Glutamina

Y / TYR – Tirosina

F / PHE – Fenilalanina

K / LYS – Lisina

L / LEU – Leucina

N / ASN – Asparagina

A / ALA – Alanina

Kcal/mol – quilocaloria por mol

CID – do inglês, *Compound ID number* (número de identificação do composto)

1. INTRODUÇÃO

O fungo de plantas neotropicais *Moniliophthora roreri* pertence à família Marasmiaceae, é o agente causador da monilíase uma doença perigosa e economicamente importante que infecta as espécies do *Theobroma cacao* e *Herrania*, apresenta uma grave ameaça aos países produtores de cacau do mundo, devido ao seu potencial devastador com danos causados exclusivamente no fruto que podem chegar a perdas de até 100% das safras de cultivo de cacau (DÍAZ-VALDERRAMA et al., 2020; EVANS, 2007; PHILLIPS-MORA, 2003; MARELLI et al., 2019; GRAMACHO et al., 2018).

Vários autores têm relatado a ocorrência da monilíase na Colômbia (1940), Venezuela (1941), Panamá (1949), Honduras (1977), Costa Rica (1978), Nicarágua (1980), Perú (1988), México (2005), Belize (2006) (PHILLIPS-MORA et al., 2006a, 2006b; EVANS, 2016), Bolívia (2012) (PHILLIPS-MORA et al., 2015), na Jamaica (2016) (JOHNSON et al., 2017) e recentemente no Brasil, na região de Cruzeiro do Sul, no Acre (junho de 2021) e com risco de disseminação para outras regiões brasileiras (BRASIL, 2021a, 2021b). Na maioria destes países produtores, existem relatos de perdas entre 30% a 70% da produção anual das safras de cacau provocados pela doença. Evans (1981) afirma que a perda da condição do Equador como produtor mundial de cacau, teve *M. roreri* como seu principal agente (DÍAZ-VADERRAMA *et al*., 2020; BAILEY *et al.*, 2018; AIME; PHILLIPS-MORA, 2005).

Porém, ainda faltam estratégias de controle de intervenção desta praga que sejam eficazes, acessíveis, e que sejam aplicáveis a pequenos produtores de cacau na maioria dos países produtores. Devido a importância econômica do cacaueiro para a região Sul do Bahia, lança-se mão de estudos envolvendo proteômica e bioinformática para encontrar potenciais alvos de estudos que sejam relevantes para o desenvolvimento do fungo *M. roreri*, e serem aplicados associados as demais medidas fitossanitárias.

Com base nessa premissa, estudos do gene ACB que codifica a proteína de ligação a acil-CoA (ACBP), mostraram que este gene esteve altamente conservado na levedura *Saccharomyces cerevisiae* fornecendo informações importantes sobre o papel biológico da ACBP e sugerindo que muitas das funções biológicas atribuídas a ele em organismos superiores podem resultar de sua capacidade de interagir com acil-CoA (ROSE, SCHULTZ; TODARO, 1992). Estudos ainda indicam que sua presença em todos os reinos, apresentando alto grau de conservação da sequência de aminoácidos de ACBP isoladas de diferentes espécies e o seu envolvimento em vários mecanismos metabólicos essenciais relacionados com o éster acil-CoA

sugerem que a ACBP é uma proteína de manutenção e o gene da ACBP tem todas as características de um gene Housekeeping (gene constitutivo) (FÆRGEMAN; KNUDSEN, 1997; MANDRUP et al, 1992; RASMUSSEN et al, 1994)

As proteínas de ligação da acil-coenzima A (ou ACBP) são conhecidas de um grande grupo de espécies de eucariotos e tem capacidade de se ligarem a um éster de acil-CoA de cadeia longa com alta afinidade; os ésteres de acil-CoA de cadeia longa são intermediários no metabolismo lipídico e modulam uma ampla gama de funções celulares importantes, como regulação do metabolismo energético, síntese lipídica, transdução de sinal, regulação do fluxo iônico e expressão gênica (MONZANI et al., 2010). ACBP é uma família multigênica (que contém mais de uma cópia no genoma) e a depender do gene de ACBP pode atuar em diferentes processos celulares. Muitas evidências experimentais *in vitro* indicam que a ACBP age no transporte e armazenamento intracelular de ésteres de acil-CoA, e trabalhos recentes indicam que ela é necessária para o tráfico de vesículas em leveduras (MANDRUP; FÆRGEMAN; KNUDSEN, 2004).

O alto grau de conservação da sequência de aminoácidos de ACBP isoladas de diferentes espécies demonstram muita semelhança, sugerindo que a estrutura desta proteína é importante para sua função fisiológica (ROSE, SCHULTZ; TODARO, 1992). A parte mais estruturalmente conservada dos ACBPs é o domínio ACB, constituído por aproximadamente 90 resíduos de aminoácidos na sua estrutura primária. Em geral, as isoformas de ACBP possuem uma estrutura na qual suas 4 α-hélices (A1-A4) acomodam o local de ligação bem exposto ao solvente em um arranjo estrutural do tipo *up-down-down-up* que é conservado em diferentes espécies (TEILUM et al., 2005). O sítio de ligação é dividido em três subsítios, classificados de acordo com a estrutura do éster acil-CoA: um para o anel de adenina, um para o grupo 3'-fosfato e um para a cadeia de carbono (KRAGELUND; KNUDSEN; POULSEN, 1999)

Segundo o estudo de MONZANI et al. (2010), a caracterização de uma nova topologia da proteína de ligação à acil-CoA (ACBP) de *Moniliophthora perniciosa* evidenciou que essa proteína é composta por 104 aminoácidos e é possivelmente uma isoforma não clássica exclusivamente de basidiomicetos, sendo expressa pela maioria das células. Além disso, os resultados sugerem que esta proteína apresenta uma afinidade crescente à ésteres de acil-CoA mais longos; ademais que, os ésteres de acil-CoA de cadeia longa são intermediários no metabolismo lipídico e modulam uma ampla gama de funções celulares importantes. Uma molécula MES ou ácido 2- (*N*- morfolino) etanossulfônico, foi encontrada no sítio de ligação

CoA vinculado a ACBP de *M. perniciosa*, localizada entre as hélices α 1 e 2, sugerindo que o sítio de ligação CoA que apresentou características diferentes de outros fungos, pode ser utilizada na triagem de inibidores de moléculas pequenas com a finalidade de propor novas moléculas com capacidade de atuarem como fungicida. A partir das perspectivas desse estudo, a ACBP de *M. perniciosa* pode ser um provável molde para construção do modelo da ACBP de *M. roreri* por serem consideradas proteínas homólogas e desse modo, podemos inferir que a ACBP de *M. roreri* seria um bom alvo para ser utilizada na triagem de inibidores.

De acordo com o estudo realizado por Fritzler, & Zhu (2011), foi possível identificar inibidores contra a proteína de ligação de acil-CoA graxo (ACBP), como potenciais novos agentes anti-*Cryptosporidium parvum*, por meio de triagem de alto rendimento de drogas conhecidas ou outras moléculas em bibliotecas de compostos, foi possível identificar 28 compostos principais mais ativos que poderiam inibir a atividade de CpACBP1, são eles: "1S, 2R-fenilpropanolamina HCl, acetazolamida, bitionol, broxiquinolina, clorpromazina, cloxacilina sódica, cloxiquina, curcumina, ácido gambógico, homatropina metilbrometo, hidralazina HCl, acetato de hidrocortisona, isoxicam, sulfossalicilato de meclociclina, cloridrato de mitoxantrona, oxacilina sódica, cloridrato de fenazopiridina, sulfato de fenelzina, fenotrina, fenitoína de sódio, pregnenolona, pristimerin, quinalizarina, rifampicina, rifaximina, desidrocolato de sódio, estreptozocina, tirotricina"; e ainda possuem uma ampla gama de bioatividades, variando de antivirais, antibacterianos, antifúngicos e antiparasitários à anti-inflamatórios e antidepressivos. A atividade anti-ACBP observada foi nova para essas drogas.

Portanto, este trabalho visou propor novas moléculas capazes de inibir a ligação entre ACBP e ésteres de acil-CoA e que reprimam o transporte de ésteres de acil-CoA e consequentemente, inibir a função desta importante proteína em *M. roreri*. Para alcançar esses objetivos uma série de metodologias moleculares *in sílico* foram empregadas como modelagem por homologia, triagem virtual baseada no receptor, acoplamento molecular (*docking*) e simulações de dinâmica molecular. Por fim, analisamos o complexo formado entre cada um dos ligantes prováveis com o receptor (MrACBP e TcACBP), para compreender a estrutura energética e dinâmica molecular dessas interações, e propor essas novas moléculas para prosseguirem com testes *in vitro* e *in vivo* no intuito de atuarem como novas perspectivas no controle da monília causada por *Moniliophthora roreri*.

2. OBJETIVOS

2.1 Objetivo Geral

Elucidar a estrutura tridimensional e propor potenciais candidatos à inibidores da proteína de ligação à acil-CoA (ACBP) de *Moniliophthora roreri*, utilizando metodologias *in sílico*.

2.2 Objetivos específicos

1) Obter as estruturas tridimensionais válidas da ACBP de *M. roreri* e *T. cacao*, através de modelagem molecular por homologia;

2) Catalogar compostos químicos naturais e sintéticos, além de drogas previamente testadas, que possuem atividade contra a proteína de ligação acil-CoA (ACBP) descritas para outros organismos.

3) Realizar a triagem de inibidores baseada na estrutura do receptor (RBVS) utilizando como fonte de busca pequenas moléculas depositadas em bancos de dados públicos de compostos químicos e testes de acoplamento molecular (*docking*).

4) Verificar se os prováveis ligantes encontrados também possuem afinidade pela ACBP de *Theobroma cacao.*

5) Descrever o comportamento da interação entre a proteína ACPB e seus compostos candidatos à inibidores, através de simulações de dinâmica molecular.

6) Selecionar pelo menos um composto candidato a inibidor da ACPB para realização de testes *in vitro* e/ou *in vivo* contra *M. roreri*.

3. REVISÃO DE LITERATURA

3.1 *Moniliophthora roreri:* descrição, nomenclatura, taxonomia e hospedeiros

Em meados dos anos de 1933, Ciferri e Parodi descreveram *Moniliophthora roreri* como o agente causador da monilíase e o nomearam de *Monilia roreri* Cif. & Par, como um ascomiceto anamórfico devido à aparente ausência de qualquer estado meiótico ou estruturas sexuais e semelhanças morfológicas com outros fitopatógenos do gênero. Entretanto, depois de um período de estudos detalhados da ultraestrutura das hifas e do processo de formação dos conídios, foi então que em 1978, reclassificado com o nome científico *Moniliophthora roreri* (Cif. & Par.) (EVANS et al., 1978). Posteriormente, devido as semelhanças em sua biologia e filogenética com o fungo *Crinipellis perniciosa* (atualmente, *Moniliophthora perniciosa*), o agente causal da vassoura de bruxa, em 2002, foi reclassificado novamente com o nome de *Crinipellis roreri* por ser o estado assexuado de *C. perniciosa* (GRIFFITH *et al.*, 2003). Por

fim, após uma caracterização molecular mais completa, por meio de testes morfológicos, citológicos e moleculares, foi transferido de volta ao gênero de basidiomiceto anamórfico, Moniliophthora HC Evans, Stalpers, Samson & Benny (EVANS et al., 1978, EVANS; HOLMES; REID, 2003; GRIFFITH et al., 2003; EVANS, 2016; PHILLIPS-MORA et al., 2007).

"A espécie *Moniliophthora roreri* ou sinônimo *Monilia roreri* Ciferri, é um fungo incluído no domínio Eukaryota; reino Fungi; filo Basidiomycota; classe Agaricomycetes; da ordem das Agaricales; e pertencente à família Marasmiaceae" (GRAMACHO et al., 2018; AIME; PHILLIPS-MORA, 2005). É o agente causal da doença monília ou monilíase do cacaueiro, uma doença economicamente importante do cacau, a distinção desta doença das demais são as características dos sintomas externos com aparência congelada ou pulverulenta dos frutos de cacau infetados (EVANS et al., 1978; CIFERRI & PARODI, 1933). Além disso, outra característica são os detalhes taxonômicos críticos de *Moniliophthora roreri* com esporogênese, os esporos são determinantes no ciclo de vida deste fungo no processo de causa da doença monilíase, ilustrados na figura 1 (EVANS, 2016), mostrando as cadeias de esporos basípetos que se encontram em diferentes estágios de seu desenvolvimento, a partir de diferenciação do esporo elipsoidal inicial na base, bem como, causando fratura na parede externa, e nas cadeias mais antigas (enrugadas) que consistem em esporos de tamanhos variáveis. Além disso, através da presença de septo doliporo em hifas vegetativas é possível identificar os fungos basidiomicetos (PHILLIPS-MORA & WILKINSON, 2007; EVANS,1981, 2007; CIFERRI & PARODI, 1933; PHILLIPS-MORA, 2003)

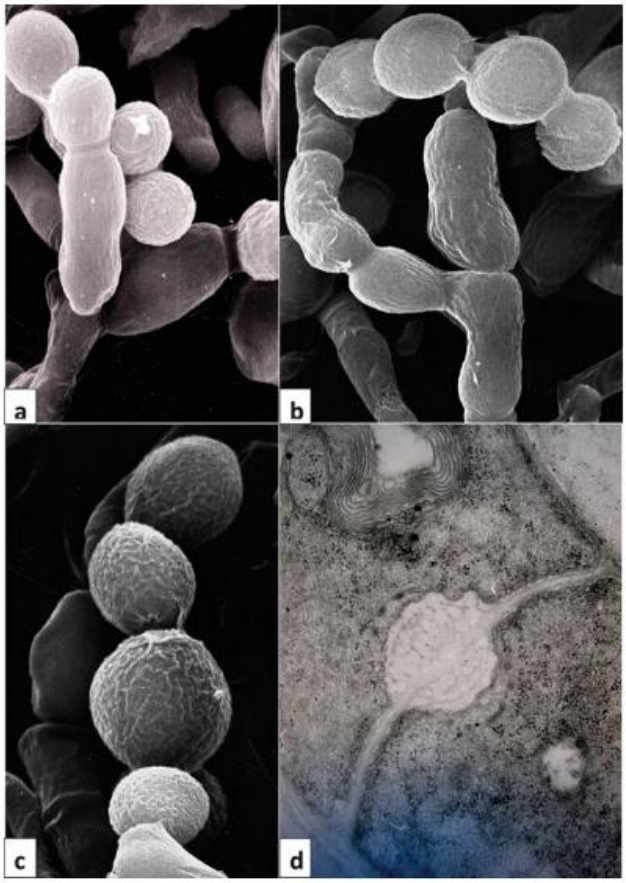

Figura 1: Microscopia eletrônica de varredura (SEM) e transmissão (TEM) para mostrar detalhes taxonômicos críticos de *Moniliophthora roreri.* (a) - (c) SEM da esporogênese, mostrando cadeias de esporos basípetos em vários estágios de desenvolvimento, a partir de (a) diferenciação do esporo elipsoidal inicial na base, e (b) fraturando a parede externa, (c) tornando-se enrugada nas cadeias mais antigas que consistem em esporos de tamanhos variáveis. (d) TEM de septo doliporo em hifas vegetativas, diagnóstico de fungos basidiomicetos. **Fonte:** EVANS (2016).

Os únicos hospedeiros conhecidos do fungo *M. roreri* pertencem as espécies dos gêneros *Theobroma* e *Herrania*, ambos gêneros dentro da família *Malvaceae.* Afetando espécies de grande importância econômica: o cacaueiro (*Theobroma cacao L.*), conhecido como a fonte do chocolate e seus derivados, e o cupuaçuzeiro [*Theobroma grandiflorum* (*Willd. ex Spreng.*) *Schum*] (GRAMACHO et al., 2018). Além de outras espécies silvestres brasileiras, presentes na Amazônia, que tem mostrado susceptibilidade ao fungo sejam em condições naturais ou até mesmo artificiais, são elas: *Theobroma bicolor*; *Theobroma subincanum; Theobroma speciosum; Theobroma obovatum; e Herrania spp*; alguns exemplos da infecção causados por *M. roreri* estão ilustrados na figura 2 (PHILLIPS-MORA; WILKINSON, 2007). Considerando a susceptibilidade aparente da maioria dos genótipos de cacau e a agressividade desse fungo ao ser transportado pelo ar, sua capacidade de sobreviver em diferentes condições ambientais e sua rápida dispersão natural e mediada pelo tráfego de pessoas, representa uma

grave ameaça substancial para as regiões produtoras de cacau em todo o mundo (AIME e PHILLIPS-MORA, 2005; EVANS, 1981).

Figura 2: Infecção por monilíase *(M. roreri)* em parentes do cacau. Sentido horário e inferior esquerdo, *Theobroma mammosum, Theobroma gileri, Theobroma bicolor, Herrania sp. e Theobroma grandiflorum*. **Fonte:** PHILLIPS-MORA; WILKINSON (2007).

3.1.2 Ciclo de vida, sintomas e características da infecção

No início do clico de vida biológico de *M. roreri*, as condições climáticas tropicais e a quantidade de esporos livres no ambiente são consideradas fatores determinantes, sendo necessário que haja a estação seca, época em que se encontram a maior quantidade de esporos disponíveis no ambiente, além de condições favoráveis de umidade.

M. roreri é um fungo hemibiotrófico, pois apresenta duas fases no ciclo de vida (figura 3), a fase biotrófica (parasítica) com presença de hifas monocarióticas e também; a fase necrotrófica (saprofítica) com hifas dicarióticas. Os sinais e sintomas da monilíase variam de acordo com a idade do fruto, mas também com o grau de severidade do ataque do patógeno. Na figura 4, mostra sintomatologia de *M. roreri* nas fases biotrófica e necrotrófica inicial; geralmente, no início apresenta sintomas na fase biotrófica, com inchaço do fruto, deformação, áreas com maturação precoce no fruto e manchas amarronzadas na cor de chocolate; e na fase necrotrófica (cerca de 60 dias após o início da infecção) se caracteriza pela união das hifas monocarióticas formando hifas dicarióticas e, posteriormente, os núcleos se fundem, aparecimento do micélio branco, que resulta num pseudoestroma na superfície do fruto infectado, consequentemente, torna-se esporulante e após isso necrótico. Por fim, ao final do

ciclo de vida biológico do fungo causador da monilíase, não há formação do corpo de frutificação e os esporos infectantes são formados a partir da meiose que ocorre dentro do núcleo diploide do micélio branco (BAILEY et al., 2013; EVANS, 1981, 2007, 2016; GRAMACHO et al, 2018; PHILLIPS-MORA, 2003; PHILLIPS-MORA & WILKINSON, 2007; MEINHARDT et al., 2014).

No ciclo de vida, representado na figura 3, os esporos produzidos em frutos necróticos infectam novos frutos, sendo que a penetração e infecção destes esporos contaminados em frutos saudáveis podem ocorrer em qualquer fase do desenvolvimento do fruto, porém são mais suscetíveis durante as primeiras fases, ou seja, quanto maior a idade dos frutos, menor será sua susceptibilidade. Após ocorrer a penetração dos esporos no fruto, ocorre com a germinação pelo meiósporos na superfície do fruto, desenvolvendo tubos germinativos monocarióticos (haploides), que penetram o fruto diretamente através da epiderme ou eventualmente, pela via dos estômatos, geralmente; ocorre um longo período de incubação (cerca de 30 dias após a infecção) intracelularmente nas células do parênquima cortical (fase biotrófica) (BAILEY *et al.*, 2013; BAILEY et al., 2018; EVANS, 1981, 2007, 2016; PHILLIPS-MORA, 2003; PHILLIPS-MORA & WILKINSON, 2007).

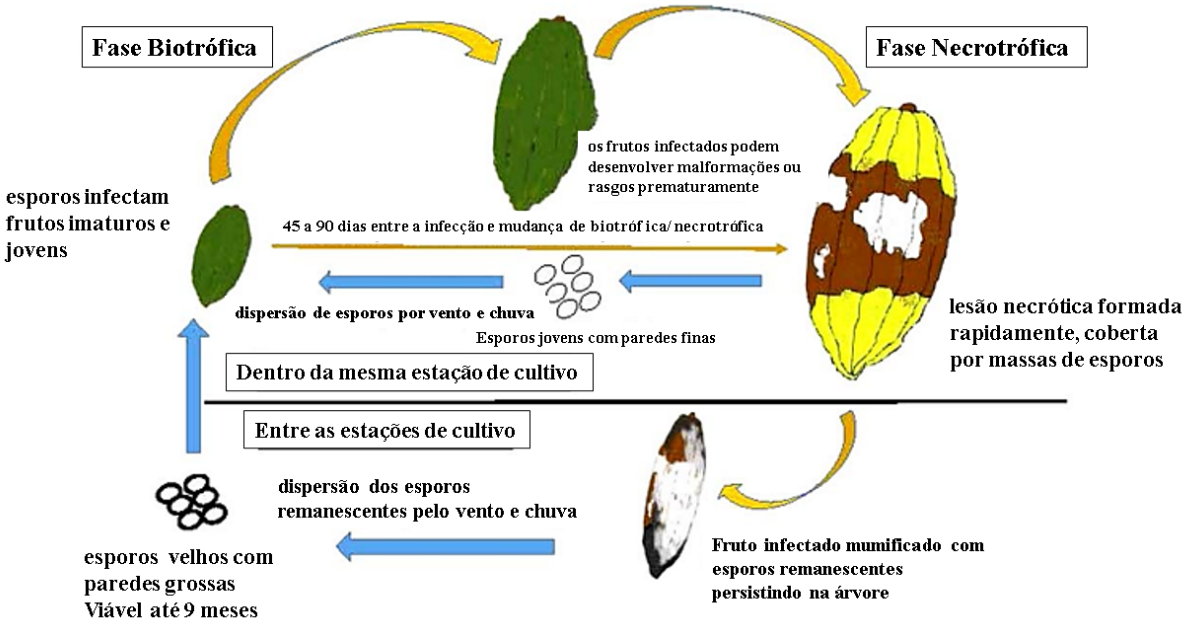

Figura 3: Ciclo da doença de *Moniliophthora roreri* em *Theobroma cacao*. Os esporos produzidos em frutos necróticos infectam novos frutos, iniciando uma fase biotrófica prolongada (45–90 dias). Assim que a fase necrotrófica é iniciada, novos esporos são rapidamente produzidos nas vagens e espalhados pelo vento e pela chuva. Os esporos remanescentes em frutos mumificados podem iniciar novas epidemias em associação com novas estações de produção. **Fonte:** Adaptado de BAILEY *et al.*, (2018).

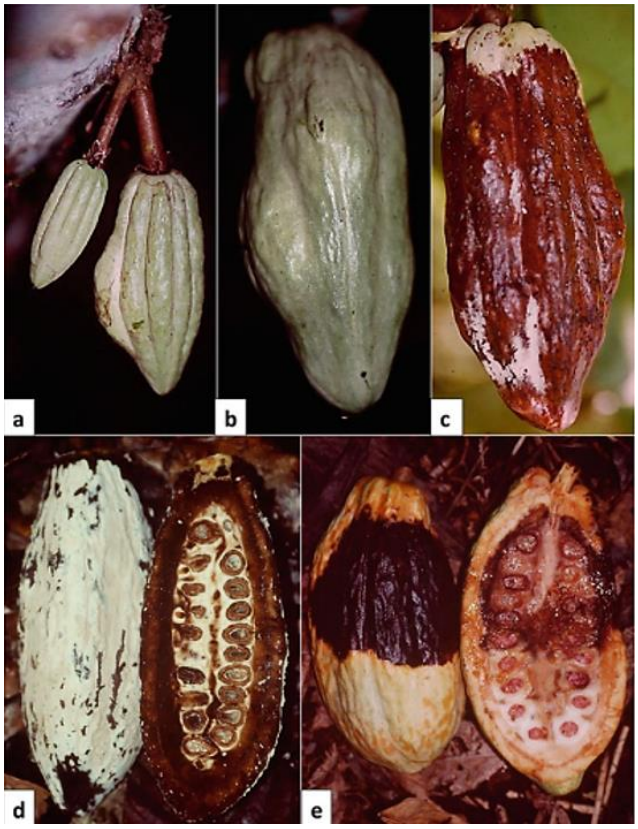

Figura 4: Sintomatologia de *Moniliophthora roreri* nas fases biotrófica e necrotrófica inicial. (a) O fruto jovem com inchaço, à direita do fruto de tamanho menor saudável. (b) Fruto verde inchado e deformado, com 2–3 meses de idade. (c) Fruto, 3–4 meses de idade, mostrando edemas e fase inicial de necrose. (d) Fruto com desenvolvimento de pseudostroma, apresenta necrose interna com compactação e destruição completa da massa do grão. (e) Lesão na cor preta externa atípica, sem o pseudostroma externo diagnóstico, mas com podridão interna úmida/aquosa; espécime da Costa Rica, onde a doença da vassoura de bruxa está ausente. **Fonte:** EVANS (2016).

3.1.3 Origem, distribuição geográfica atual e impacto econômico

Durante muitos anos, o centro de origem do *M. roreri* foi considerado o Equador, devido aos primeiros relatos autênticos sobre a origem do patógeno no final de 1917, quando J.B. Rorer esteve na região de Quevedo, Los Ríos, no Equador para identificar o agente causador de uma doença que deteriorou gravemente as safras de cacau daquela região (RORER, 1918). Rorer enviou várias amostras coletadas dos frutos de cacau infectados e doentes para o Dr. R.E. Smith, da Universidade da Califórnia, que determinou como sendo um fungo o causador da doença, que ele denominou de *Monilia sp*. (PHILLIPS-MORA, 2003; SANCHEZ-MORA; FIALLOS, 2012). Entre os anos de 1917 e 1925, Rorer fez novamente várias visitas ao Equador para avaliar os danos causados e testar o controle químico para esta doença. O Equador foi considerado por muitos autores como o centro de origem do patógeno e o fungo foi considerado endêmico nesta região, principalmente no *Theobroma bicolo*r (EVANS et al., 1981; PHILLIPS-MORA, 2003; EVANS, 2016). Consequentemente, o micologista Rorer mais

tarde se tornou reconhecido como sendo o pioneiro na pesquisa da doença (PHILLIPS-MORA, 2003; BAILEY et al., 2018).

Segundo Phillips-Mora (2003) e algumas referências encontrada durante a pesquisa bibliográfica, sugerem fortemente que a doença causada por *M. roreri*, pode ter aparecido pela primeira vez em 1817 e sustentam com o registro dos surtos epidêmicos da doença no departamento do Norte de Santander, na Colômbia, localidade onde foi encontrada a maior concentração de diversidade genética do patógeno, os sintomas possivelmente representam o da monilíase, ocorrendo cerca de 100 anos antes do caso relatado por Rorer (1918) no Equador. Além disso, evidências moleculares mediante os polimorfismos de comprimentos de fragmentos amplificados (AFLP), perfis ISSR e dados de sequência ITS, indicam que existe grande diversidade genética de *M. roreri* na Colômbia; também, evidenciando que nesta região pode ter sido a provável origem da doença, em vez do Equador (EVANS, 2002; PHILLIPS-MORA; WIKINSON, 2007; PHILLIPS-MORA; AIME; WILKINSON, 2007).

A monilíase encontra-se restrita ao continente Americano, estando presente em todos os países produtores de cacau da América Tropical devidos as condições climáticas tropicais, onde foi relatada por muitos pesquisadores ao longo de um período de quase 200 anos, isolada geograficamente no noroeste da América do Sul. Porém, aumentou os seus limites geográficos, como representado na figura 5, para as regiões da Américas do Norte (México), do Sul, Central (EVANS, 2016) e Caribe (JOHNSON et al., 2017). Foi relatada na Colômbia (1817) (PHILLIPS-MORA, 2003), Venezuela (1941), Panamá (1949), Honduras (1977), Costa Rica (1978), Nicarágua (1980), Perú (1988), México (2005), Belize (2006) (PHILLIPS-MORA et al., 2006a, PHILLIPS-MORA et al., 2006b; EVANS, 2016), Bolívia (2012) (PHILLIPS-MORA et al., 2015), na Jamaica (2016) (JOHNSON et al., 2017) e recentemente no Brasil, na região de Cruzeiro do Sul, no Acre (junho de 2021) (BRASIL, 2021a, 2021b).

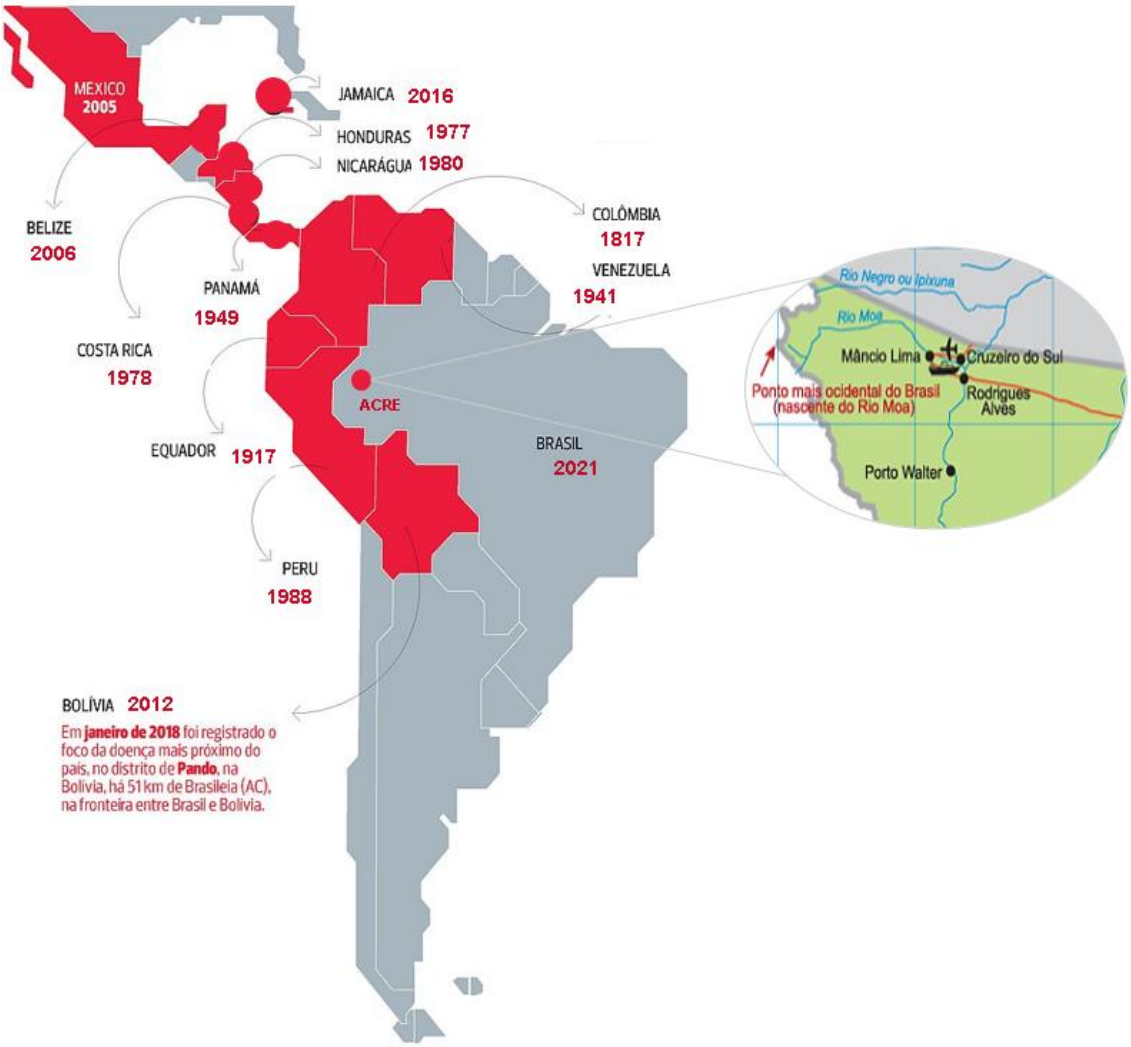

Figura 5: Distribuição geográfica da Monilíase do cacau causada por *Moniliophthora roreri.* Fonte: Adaptado da Infografia de MIRANDA, Morgana (2018); ACRE, Mapas.

A monilíase causada por *M. roreri* têm sido extremamente devastadora nas culturas de cacau chegando as perdas dos frutos pelos danos causados acima de 30%, que podem exceder à 90%, representando uma grave ameaça e um dos principais fatores limitantes da produção em muitos países como Equador, Costa Rica, Peru, Colômbia, onde sua introdução foi relatada devido ao abandono total do cultivo de cacau (AIME; PHILLIPS-MORA, 2005; PHILLIPS-MORA, 2003, PHILLIPS-MORA & WILKINSON, 2007; EVANS, 1981, 2002) . A indicação de medidas fitossanitárias por órgãos do governo brasileiro, bem como análises de risco do potencial econômico sobre a probabilidade de introdução e dispersão da monilíase do cacaueiro no Brasil já tinham sidos relatados por muitos pesquisadores, considerando a proximidade dos focos de infecção por *M. roreri*, nas regiões fronteiriças com o Brasil que seria a principal rota de entrada da praga no país (NASCIMENTO, 2014; SGRILLO, 2010, GRAMACHO et al., 2018)

Entretanto, foi constatado a primeira detecção de introdução do foco da praga *M. roreri*, conhecida como Monilíase do Cacaueiro, em área residencial urbana no município de Cruzeiro do Sul, Estado do Acre no Brasil, destacado na figura 5, foi confirmado pelo Ministério da Agricultura Pecuária e Abastecimento (MAPA) em 08 de julho de 2021. A partir desta constatação, foi realizado alinhamento de ações imediatas de controle, contenção, erradicação e prevenção da disseminação desta praga quarentenária no Brasil, para mistigar o risco de contaminação da lavoura cacaueiro do estado e bem como, sua dispersão para outras regiões e estados brasileiros, dentre as principais medidas fitossanitárias incluem a fiscalização dos plantios nesta área, orientação dos produtores quanto às medidas preventivas que podem ser adotadas, utilizado do plano de contingência da monilíase, realizados pela equipe do MAPA em conjunto com as demais instituições oficiais de Sanidade Vegetal e de pesquisa, do governo federal brasileiro (BRASIL, 2021a, 2021b, 2021c).

A valoração do impacto econômico proposto por Nascimento (2014) com a introdução de *M. roreri* no Brasil, e a interação deste patógeno com os hospedeiros e com o ambiente, de acordo com os prejuízos estimados poderia causar um impacto de aproximadamente 100 milhões de dólares a cada ano, contabilizando-se somente os impactos diretos ao produtor rural, caso medidas de controle eficazes e contenção não sejam eficazes, portanto, medidas fitossanitárias e grupos de cooperação entre os governos e outros órgãos tem se reunido para planejar estratégias e propor medidas que possam prevenir essa disseminação para outros estados (BAHIA, 2021a, 2021b; GRAMACHO et al., 2018)

A produção de cacau e chocolate é a principal atividade econômica em mais de 100 municípios baianos, gerando emprego, renda e divisas para o Estado da Bahia, considerando a possibilidade de disseminação do patógeno para outros estados ainda indenes podendo vir a ocorrer principalmente pelo trânsito de pessoas, de material vegetal, de artigos regulamentados e pelo trânsito postal; o auto potencial de dispersão pode causar perdas de até 100% na produção e provocando graves prejuízos, com reflexos econômicos, sociais e ambientais (BAHIA, 2021a, 2021b; GRAMACHO et al., 2018)

3.1.4 Principais estratégias de controle no manejo integrado

No momento, não se dispõe de técnicas que sejam totalmente eficazes para o seu controle, sendo o manejo integrado a forma mais eficiente de controle simultâneo. O manejo integrado de pragas consiste no agrupamento de princípios básicos: aspectos ecológicos, econômicos e também, em teorias sociológicas; que no contexto do meio ambiente está

associado a dinâmica populacional do patógeno; utilizando todas as técnicas e métodos que sejam adequados e compatíveis de tal modo, que permita manter as populações das pragas e patógenos em níveis inferiores ao que causam potenciais danos econômicos; devendo estes serem economicamente viáveis e ambientalmente aceitáveis (SUÁREZ; HERNÁNDEZ, 2010; TIRADO-GALLEGO; LOPERA-ÁLVAREZ; RÍOS-OSORIO, 2016; BRASIL, 2021c; BAHIA, 2021a, 2021b).

Aplicando o manejo integrado como método de controle da monilíase do cacaueiro, envolve o controle cultural, que consiste na remoção de frutos infectados antes que ocorra a esporulação, deixando-os sobre a superfície do solo; realiza-se ainda a drenagem do solo a fim de reduzir a umidade na área; redução na forma de sombreamento; o controle de ervas daninhas; a poda de rebaixamento das copas dos cacaueiros que tem finalidade facilitar a remoção de frutos infectados, o controle químico, através de aplicações de fungicidas, são eles: chlorothalonil (além de reduzir a incidência, pode aumentar a produtividade da cacauicultura), utilização do óxido cuproso, ou associação dos dois simultaneamente; o controle biológico associado com o controle genético, consiste na utilização de genótipos resistentes do cacau à monilíase, que tem sido o principal foco de muitos programas de pesquisa onde a doença provocou grande desastres econômicos e ambientais, especialmente no Equador, Costa Rica e Colômbia (CRONSHAW, 1979; SUÁREZ; HERNÁNDEZ, 2010; SANCHEZ-MORA & FIALLOS, 2012; EVANS, 2016;TIRADO-GALLEGO; LOPERA-ÁLVAREZ; RÍOS-OSORIO, 2016)

Além disso, considerando o risco da introdução, e disseminação para outros estados do Brasil após a constatação da detecção do foco de monilíase na cidade de Cruzeiro do Sul, estado do Acre, sobre a coordenação do Ministério da Agricultura Pecuária e Abastecimento (MAPA) emitiu a Instrução Normativa nº 112, registrada no Diário Oficial, publicada em 15 de dezembro de 2020, que institui o plano nacional de prevenção e vigilância de *Moniliophthora roreri* – PNPV/Monilíase, com o objetivo de detalhar as medidas fitossanitárias para prevenção e controle, supressão e erradicação da monilíase, propondo diretrizes e os procedimentos operacionais para aplicação dessas medidas por meio de: educação fitossanitária, capacitação, elaboração, coordenação e execução de ações para prevenção e para contingência em casos de suspeita do foco de *M. roreri*.

Nos últimos anos, devido às limitações das estratégias de contingência para o controle da monilíase e o risco de disseminação, a utilização de ferramentas computacionais associadas

à pesquisa genômica e proteômica estão entre as principais estratégias adotadas para serem associadas as demais medidas fitossanitárias.

3.2 Uma visão geral da estrutura e função de proteínas

As proteínas são as macromoléculas mais abundantes e que participam de processos celulares vitais nos seres vivos, por serem consideradas instrumentos moleculares através do qual a informação genética é expressa, e podem atuar numa ampla diversidade de funções estruturais e dinâmicas importantes como regulação, catalisadores de reações químicas, transporte de moléculas, digestão, defesa do organismo, estruturas de sustentação de células e tecidos, e outras.

As proteínas são formadas por polipeptídeos (constituídos por muitos aminoácidos unidos através de ligações peptídicas) e de acordo com as sequências de aminoácidos que compõem as proteínas e os ângulos que esses aminoácidos assumem, podem assumir várias conformações tridimensionais. São divididas em quatro níveis estruturais (figura 6): (1) a estrutura primária é formada pela sequência de aminoácidos unidos por ligações peptídicas e ligações dissulfeto; (2) a estrutura secundária que corresponde a regiões localizadas da estrutura proteica resultante de ligações de hidrogênio, como padrões de dobramento da cadeia principal e as principais estruturas secundárias são as α-hélice (estrutura helicoidal) e outras proteínas podem apresentar também a folha-β pregueada (estrutura em forma de "zigue-zague" e se assemelha a uma folha dobrada), esta última pode apresentar as fitas dispostas paralelamente umas às outras (N-terminal e o C-terminal são correspondentes), ou antiparalelas apontando em sentidos/direções opostas (o N-terminal de uma fita se aproxima do C-terminal da outra fita), e regiões de Volta-β que conectam estas estruturas secundárias, geralmente encontradas na área externa da proteína e seus aminoácidos não formam pontes de hidrogênio (NELSON & COX, 2014).

Na figura 6, (3) a estrutura terciária se refere a estrutura tridimensional de um polipeptídeo, que surge das interações entre os grupos R (ligações de hidrogênio, ligações iônicas, forças de dispersão London, interações dipolo-dipolo) dos aminoácidos que a compõem, como também são constituídas de interações hidrofóbicas em sua estrutura terciária, que ocorre devido a junção de aminoácido não polares, e grupos hidrofóbicos R se concentram no interior da proteína, e os aminoácidos hidrofílicos interagem com as moléculas de água no exterior; e a (4) estrutura quaternária que é formada pela junção de várias cadeias polipeptídicas

(subunidades) ou estruturas terciárias, onde cada uma delas pode ser afetadas por outras estruturas que interagem com estas subunidades e consequentemente afetar sua subfunção, além disso, possui os mesmos tipos de interações da estrutura terciária, como também apresentam as interações fracas (ligações de hidrogênio e forças de dispersão de London) (NELSON & COX, 2014)

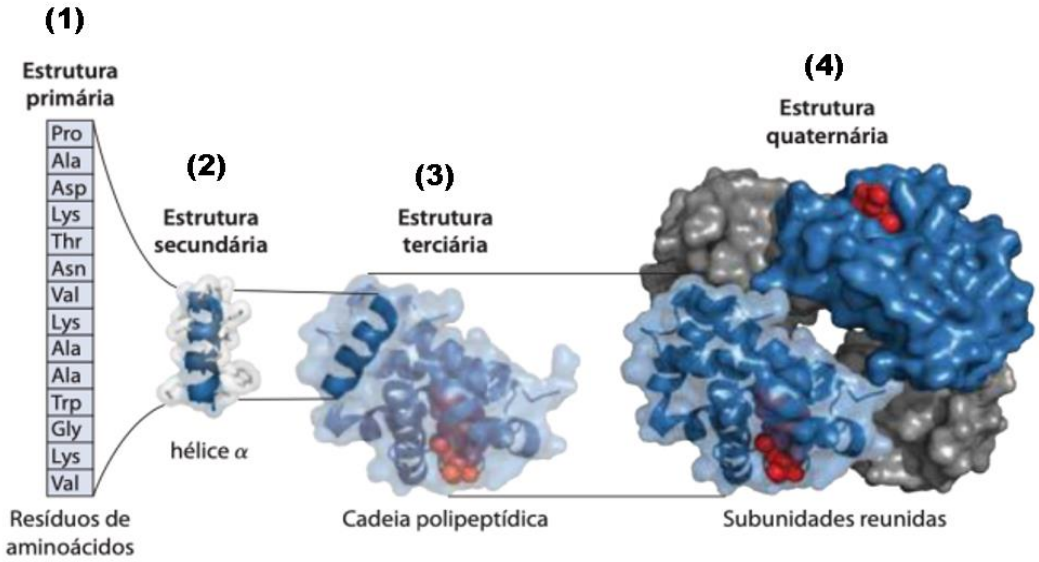

Figura 6: Níveis de estrutura nas proteínas. (1) Uma sequência de aminoácidos unidos por ligações peptídicas e ligações dissulfeto formam a estrutura primária. (2) O polipeptídeo resultante pode ser disposto em unidades de estrutura secundária, nesse exemplo a estrutura secundária se apresenta como α-hélice. A hélice é uma parte que constitui a (3) estrutura terciária do polipeptídeo dobrado, e a (4) estrutura quaternária é composta pelas subunidades da estrutura terciária (proteína com multissubunidades), exemplificado na hemoglobina. **Fonte:** Adaptado de NELSON & COX (2014).

É importante ressaltar que proteínas grandes (com mais de 100 resíduos de aminoácidos) normalmente se dobram em subunidades, que são conhecidas como domínios proteicos, é a região da proteína de maior densidade e que possui mais dobras, os quais podem conter alto nível de interações, e cada domínio apresenta uma determinada função, podendo se ligar a pequenas moléculas ou interagir com outras proteínas, ou pode ser estrutural se referindo a um componente estável da estrutura proteica. Os domínios homólogos em proteínas possuem uma estrutura conservada, com similaridades entre as sequências de aminoácidos e funções comuns. Uma proteína pode ser formada por uma única cadeia polipeptídica ou por duas ou até mais cadeias polipeptídicas contendo somente um domínio compartilhado entre elas ou mais domínios em cada uma das cadeias polipeptídicas. Além disso, ainda possuem motivos (também conhecidos como *fingerprint* ou impressão digital) que são regiões curtas e altamente conservadas nas sequências de proteínas (NELSON & COX, 2014)

Portanto, uma prévia compreensão dos conceitos básicos das estruturas que determinam as funções biológicas importantes em proteínas é fundamental para o estudo da relação de suas estruturas na determinação da função biológica, principalmente em organismos fitopatogênicos, e na modelagem dessas estruturas proteicas por técnicas de bioinformática estrutural.

3.3 Proteína de ligação à acil-CoA (ACBP)

A proteína de ligação acil-CoA (ACBP) é uma proteína citosólica pequena (aproximadamente10 kD) capaz de se ligar com alta afinidade e especificidade aos ésteres de acil-CoA de cadeia média e longa e transportá-los para dentro da célula (ROSE, SCHULTZ; TODARO, 1992; FÆRGEMAN; KNUDSEN, 1997). Os ésteres de acil-CoA de cadeia longa são utilizados como intermediários fundamentais na síntese de lipídeos e também, na degradação dos ácidos graxos; além dessa função basal, evidências indicam que ésteres de acil-CoA de cadeia longa também têm uma função importante na regulação do metabolismo intermediário de lipídeos, na regulação gênica e como moléculas de sinal; trabalhos recentes indicam ainda que a ACBP é necessária para o tráfico de vesículas em leveduras. Portanto, a expressão da ACBP em todas as células e tecidos sugerem que sua função está associada a uma ou mais funções bioquímicas essenciais na célula (ROSE, SCHULTZ; TODARO, 1992; FÆRGEMAN; KNUDSEN, 1997; MANDRUP; FAERGEMAN; KNUDSEN, 2004).

O gene ACB que codifica a proteína de ligação a acil-CoA (ACBP), esteve altamente conservado na levedura *Saccharomyces cerevisiae* fornecendo informações importantes sobre o papel biológico da ACBP e sugerindo que muitas das funções biológicas atribuídas a ele em organismos superiores podem resultar de sua capacidade de interagir com acil-CoA. A proteína ACBP foi denominada inicialmente como inibidor de ligação do Diazepam (DBI) ou Endozepina (EP), com base em sua capacidade de inibir a ligação do Diazepam ao receptor GABA, baseado em análises da sequência de aminoácidos de proteínas isoladas de humanos, bovinos e ratos que demonstraram que DBI e EP eram produto do mesmo gene ACB (ROSE, SCHULTZ; TODARO, 1992).

A família multigênica dentro da superfamília ACBP é altamente conservada em diferentes espécies de plantas e eucariotos. Os homólogos de ACBP encontram-se altamente conservados presente em todos os organismos distribuídos nos quatro reinos eucarióticos, *Animalia*, *Plantae*, *Fungi* e *Protista*, e em algumas espécies de eubactérias com alta identidade entre as sequências, além disso, o número de aminoácidos varia de acordo com

a espécie, como é ilustrado na figura 7; desse modo, sua presença em todos os reinos e o seu envolvimento em vários mecanismos metabólicos essenciais relacionados com o éster acil-CoA sugerem que a ACBP é uma proteína de manutenção e o gene da ACBP tem todas as características de um gene *housekeeping* (gene de manutenção) (FÆRGEMAN; KNUDSEN, 1997; MANDRUP et al, 1992; RASMUSSEN et al, 1994). Desse modo, o alto grau de conservação da sequência de aminoácidos de ACBP isoladas de diferentes espécies demonstram muita semelhança, indicando que os homólogos da ACBP derivam de um ancestral comum e sugerindo que a estrutura desta proteína é importante para sua função fisiológica, além disso desempenhando funções semelhantes nesses organismos (ROSE, SCHULTZ; TODARO, 1992).

```
Human-1     WGDLWLLPPASANQGTGTEAEFEKAAEEVRHLKTKPSDEEMLFIYGHYKQATVGDIN-TER   60
Human-2     -----------------SQAEFEKAAEEVRHLKTKPSDEEMLFIYGHYKQATVGDIN-TER   43
Rat         -----------------SQADFDKAAEEVKRLKTQPTDEEMLFIYSHFKQATVGDVN-TDR   43
Mouse       -----------------SQADFDKAAEEVKRLKTQPTDEEMLFIYSHFKQATVGDVN-TDR   43
Bovine      -----------------SQAEFDKAAEEVKHLKTKPADEEMLFIYSHYKQATVGDIN-TER   43
Pig         -----------------SQAEFEKAAEEVKNLKTKPADDEMLFIYSHYKQATVGDIN-TER   43
Dog         -----------------SQAEFDKAAEDVKHLKTKPADDEMLYIYSHYKQATVGDIN-TER   43
Tortoise    -----------------SQAEFDKAAEEVKQLKSQPTDEEMLYIYSHFKQATVGDIN-TER   43
Duck        -----------------AEAAFQKAAEEVKQLKSQPSDQEMLDVYSHYKQATVGDVN-TDR   43
Chicken     -----------------SEAAFQKAAEEVKELKSQPTDQEMLDVYSHYKQATVGDVN-TDR   43
Frog        ---------------MSPQADFDKAAGDVKKLKTKPTDDELKELYGLYKQSTVGDINIEC-   45
M. secta    ---------------MSLQAQFDQAASNVRNLKSLPSDNDLLELYALFKQASAGDADPANR   46
Drosophila  ---------------MVSE-QFNAAAEKVKSLTKRPSDDEFLQLYALFKQASVGDND-TAK   43
Yeast-1     ----------------VSQ-LFEEKAKAVNELPTKPSTDELLELYALYKQATVGDND-KEK   43
Yeast-2     ----------------VSQ-LFEEKAKAVNELPTKPSTDELLELYGLYKQATVGDND-KEK   43
B. napus    ---------------MGLKEDFEEHAEKVKKLTASPSNEDLLILYGLYKQATVGPVT-TSR   45
A. thaliana ---------------MGLKEEFEEHAEKVNTLTELPSNEDLLILYGLYKQAKFGPVD-TSP   45
Cotton      ---------------MGLKEEFEEHAEKVKTLPAAPSNDDMLILYGLYKQATVGPVNT-SR   44
ELP         ----------------MSQVEFEMACASLKQLKGPVSDQEKLLVYSFYKQATQGDCNIP-V   44

Human-1     ---------------PGMLDFTGKAKWDAWNELKGTSKEDAMKAYINKVEELKKKYGI-    103
Human-2     ---------------PGMLDFTGKAKWDAWNELKGTSKEDAMKAYINKVEELKKKYGI-     86
Rat         ---------------PGLLDLKGKAKWDSWNKLKGTSKENAMKTYVEKVEELKKKYGI-     86
Mouse       ---------------PGLLDLKGKAKWDSWNKLKGTSKESAMKTYVEKVDELKKKYGI-     86
Bovine      ---------------PGMLDFKGKAKWDAWNELKGTSKEDAMKAYIDKVDELKKKYGI-     86
Pig         ---------------PGILDLKGKAKWDAWNGLKGTSKEDAMKAYINKVDELKKKYGI-     86
Dog         ---------------PGLLDLRGKAKWDAWNQLKGTSKEDAMKAYVNKVDELKKKYGI-     86
Tortoise    ---------------PGFLDFKGKAKWDAWDALKGMAKEEAMKAYIAKVEELKGKYGI-     86
Duck        ---------------PGMLDFKGKAKWDAWNALKGMSKEDAMKAYVAKVEELKGKYGI-     86
Chicken     ---------------PGMLDFKGKAKWDAWNALKGMSKEDAMKAYVAKVEELKGKYGI-     86
Frog        ---------------PGMLDLKGKAKWDAWNLKKGLSKEDAMSAYVSKAHELIEKYGL-     88
M. secta    ---------------PGLLDLKGKAKFDAWHKKAGLSKEDAQKAYIAKVESLIASLGLQ     90
Drosophila  ---------------PGLLDLKGKAKWEAWNKQKGKSSEAAQQEYITFVEGLVAKYA--     86
Yeast-1     ---------------PGIFNMKDRYKWEAWENLKGKSQEDAEKEYIALVDQLIAKYSS-     86
Yeast-2     ---------------PGIFNMKDRYKWEAWENLKGKSQEDAEKEYIAYVDNLIAKYSS-     86
B. napus    ---------------PGMFSMKERAKWDAWKAVEGKSTDEAMSDYITKVKQLLEAEASSASA  92
A. thaliana ---------------PGMFSMKERAKWDAWKAVEGKSSEEAMNDYITKVKQLLEVAASKAST  92
Cotton      ---------------PGMFNMREKYKWDAWKAVEGKSKEEAMGDYITKVKQLFEAAGSS     89
ELP         ---------------PPATDVRAKAKYEAWMVNKGMSKMDAMRIYIAKVEELKKKEPC-     87
```

Figura 7: Comparação de sequências de aminoácidos de ACBPs de espécies diferentes. Em destaque (negrito) os resíduos idênticos entre as sequências. **Fonte:** FÆRGEMAN & KNUDSEN, (1997).

O domínio ACB comum para todos os ACBPs é a parte mais estruturalmente conservada dos ACBPs. A estrutura da ACBP se apresenta em um arranjo estrutural do tipo *up-*

down-down-up que é conservado em diferentes espécies (GUO et al., 2017; TEILUM et al., 2005). Nesta conformação, mostrada (figura 8) na OsACBP1 de *Oryza sativa*, a hélice 1 (H1) interage fortemente com a hélice 2 (H2), com a qual é conectada por uma pequena volta, ou seja, unidas por loops designado como loop H1-H2 e as duas estão dispostas na orientação antiparalela entre si, assim como hélice 3 mais curta (H3) e hélice 4 mais longa (H4), mas a hélice 2 e 3 são orientadas paralelamente e conectadas por uma volta de resíduos (loop H2-H3). A H3 é desarticulada das hélices H1 e H4, com região de loop entre H3-H4, resultando em apenas quatro interfaces de interação hélice-hélice, como é comumente encontrado em estruturas de quatro hélices. Estas características confere a ACBP um enovelamento único entre as conformações conhecidas de quatro hélices (H1-H4), portanto é provável que a arquitetura do feixe de quatro hélices seja conservada em outras estruturas de ACBP determinadas até o momento, especialmente para ACBPs com menos de 125 resíduos de aminoácidos; porém existem raras exceções, como a ACBP característico de *M. perniciosa* que contém uma quinta hélice adicional no terminal C (TEILUM et al, 2005; KRAGELUND; KNUDSEN; POULSEN, 1999; GUO et al., 2017; MONZANI et al., 2010).

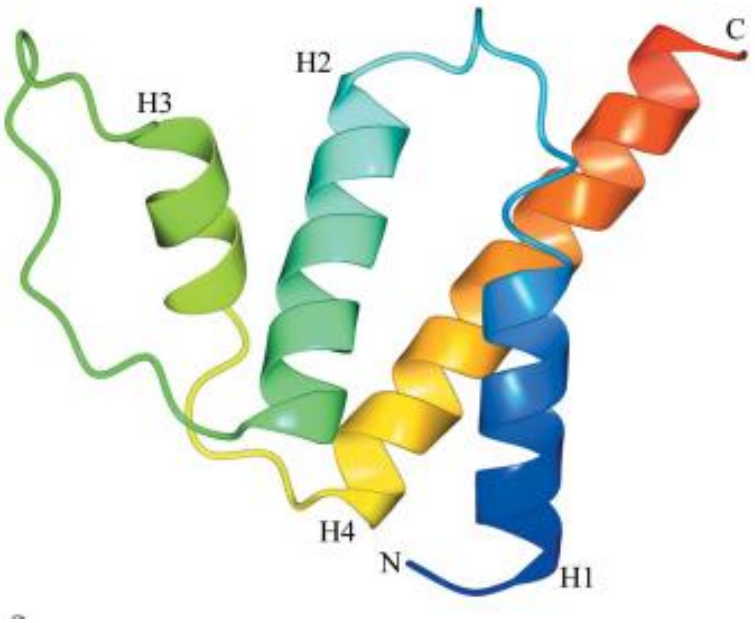

Figura 8: Estrutura de OsACBP1 (*Oryza sativa*) na representação de fita colorida. Usando uma transição gradual de azul no terminal N (marcado como N) para vermelho no terminal C (marcado como C). As quatro hélices (H1 – H4) são rotuladas em seus terminais N. **Fonte:** GUO et al., (2017).

Um estudo realizado por Monzani e colaboradores (2010) revelou uma nova topologia para a proteína de ligação à acil-CoA (ACBP) de *Moniliophthora perniciosa* ilustrado na figura 9, composto por 104 aminoácidos e é possivelmente uma isoforma não clássica exclusivamente

de basidiomicetos. Seus resultados sugerem uma afinidade crescente de ACBP para ésteres de acil-CoA mais longos, como miristoil-CoA para araquidoil-CoA, e a modelagem de melhor ajuste indicou dois sítios de ligação. Os ésteres de acil-CoA de cadeia longa são intermediários no metabolismo lipídico e modulam uma ampla gama de funções celulares importantes, como regulação do metabolismo energético, síntese lipídica, transdução de sinal, regulação do fluxo iônico e expressão gênica. ACBP passou por uma mudança de um estado monomérico que liga a lauroil-CoA e miristoil-CoA preferencialmente para um dimérico, principalmente quando ligado a ésteres de acil-CoA de cadeia longa. A estrutura da proteína foi determinada em resolução de 1,6 Å demonstrou que a nova topologia para ACBP, contendo em sua estrutura 3D, regiões com cinco α-hélices em vez de quatro, α-hélices 1, 2, 3 e 4 adotaram um arranjo agrupado que é único das dobras de quatro hélices previamente determinadas de ACBP, enquanto α-hélices 1, 2, 4 e 5 formaram um pacote clássico de quatro hélices (Figura 9 A, B).

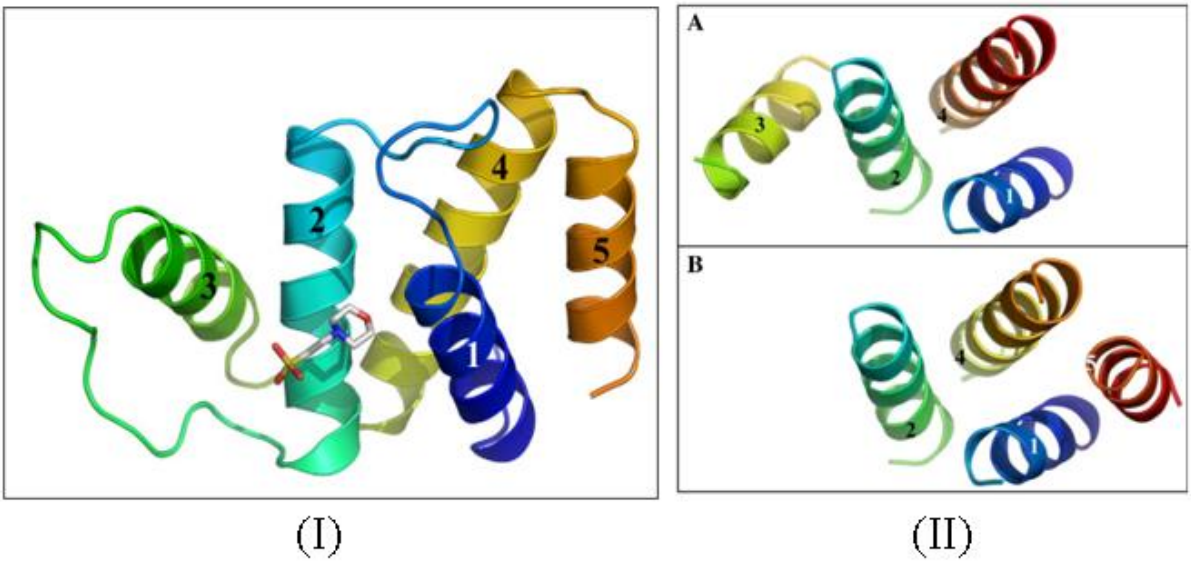

Figura 9: Estrutura da ACBP em *Moniliophthora perniciosa*. (I) Os N-terminais são mostrados em azul e os C-terminais são mostrados em laranja, com uma hélice α extra (5). Uma molécula MES localizado no local de ligação do CoA. II. Empacotamento das α-hélices, (A) topologia típica de ACBP é um arranjo único entre quatro dobras em hélice conhecidas, formadas pelas hélices 1, 2, 3 e 4; (B) um feixe de quatro hélices clássico incluindo as hélices 1, 2, 3 e 5. A quinta hélice é única para ACBP de *M. perniciosa*. **Fonte:** Adaptado de MONZANI et al., (2010).

Os modos de ligação do ligante entre as ACBPs no sítio ativo são variados. Na figura 10, uma molécula MES ou ácido 2- (*N*- morfolino) etanossulfônico (usado como tampão na cristalização), foi encontrada no sítio de ligação CoA vinculado a ACBP de *M. perniciosa*, localizada entre as hélices α 1 e 2, sugerindo que o sítio ativo de ligação CoA que contém os aminoácidos: Q14 (glutamina), Y33 (tirosina), F36 (fenilalanina), K37 (lisina), K59 (lisina) e Y78 (tirosina); pode ser utilizado na triagem de inibidores de moléculas pequenas com a

finalidade de propor novas moléculas com capacidade de atuarem como fungicida (MONZANI et al., 2010). Portanto, essa isoforma exclusivamente de basidiomicetos com características diferentes de outros fungos, a ACBP de *M. perniciosa* é específica da espécie e pode ser considerada um potencial molde para a construção do modelo da ACBP de *M. roreri*.

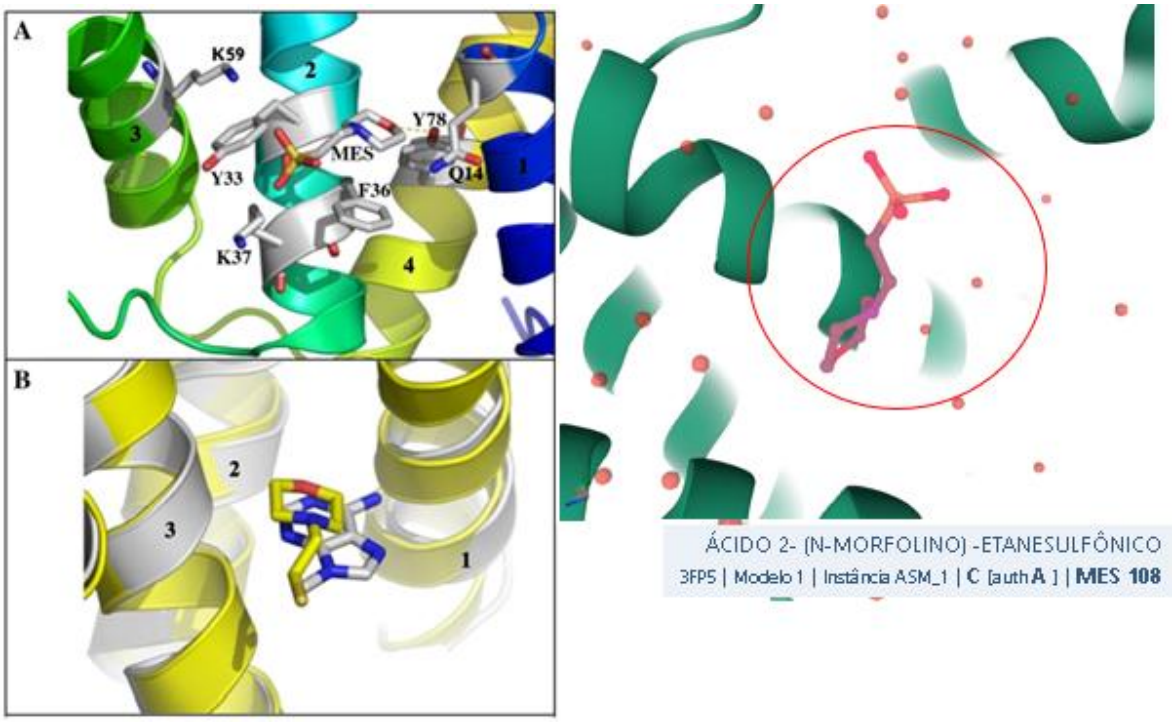

Figura 10: MES vinculado a *M. perniciosa* ACBP. (A) Resíduos de aminoácidos envolvidos na ligação de MES a ACBP. (B) MES sem um grupo fosfato é ligado no mesmo local que a adenina de CoA, mostrado por sobreposição com o ACBP humano ligado (RCSB PDB 2CB8) (C) MES (RCSB PDB 3FP5), adaptado. **Fonte:** Adaptado de MONZANI et al., (2010); RCSB PDB 3FP5.

3.3.1 Inibidores da ACBP: novas drogas com atividade anti-ACBP

Fritzler e Zhu (2011) realizaram um estudo para identificar inibidores contra a proteína de ligação de acil-CoA graxo (ACBP) de *Cryptosporidium parvum*, causador de uma infecção oportunista em pacientes com AIDS e sem tratamento eficaz disponível, por meio de triagem de alto rendimento de drogas conhecidas ou outras moléculas em bibliotecas de compostos, como potenciais novos agentes antiparasitários. Neste estudo foram testadas *in vitro* a potência e eficácia das moléculas candidatas em relação ao crescimento do parasita, além do ensaio de citotoxicidade *in vitro*, 28 compostos (Tabela 1) principais mais ativos que poderiam inibir a atividade de CpACBP1; e ainda possuem uma ampla gama de bioatividades, variando de antivirais, antibacterianos, antifúngicos e antiparasitários a anti-inflamatórios e antidepressivos. A atividade anti-ACBP observada foi nova para essas drogas.

Tabela 1 - Os 28 principais compostos para inibição da ligação da proteína de ligação acil Coa graxo (CpACBP1) para NBD-palmitoil-CoA, e as bioatividades principais. Em "cinza" foram os compostos (broxiquinolina, cloxiquina, cloxacilina – que possuem atividades antibacterianas e antifúngicas - e, desidrocolato de sódio – semi-sal biliar sintético e agente colerético) que exibiram atividades anti-criptosporídicas, ou seja, os quatro compostos podem inibir o crescimento do parasita *in vitro*.

Nomes dos compostos	Bioatividade principal
1S, 2R-fenilpropanolamina HCl	descongestionante, anoréxico
acetazolamida	inibidor da anidrase carbônica, diurético, anti-glaucoma
bitionol	anti-helmíntico, anti-séptico
broxiquinolina	**anti-infeccioso, desinfetante**
clorpromazina	antiemético, antipsicótico
cloxacilina sódica	**antibacteriano**
cloxiquina	**antibacteriano, antifúngico**
curcumina	antibacteriano, antifúngico, inibidor de lipo/ciclooxigenase
ácido gambógico	antiinflamatório, citotóxico, inibe células HeLa *in vitro*
homatropina metilbrometo	anticolinérgico (oftálmico)
hidralazina HCl	anti-hipertensivo
acetato de hidrocortisona	glicocorticóide, antiinflamatório
isoxicam	anti-inflamatório
sulfossalicilato de meclociclina	antibacteriano
cloridrato de mitoxantrona	antineoplásico
oxacilina sódica	antibacteriano
cloridrato de fenazopiridina	analgésico
sulfato de fenelzina	antidepressivo
fenotrina	ectparasiticida
fenitoína de sódio	anticonvulsivante, antiepiléptico
pregnenolona	glicocorticóide, antiinflamatório
pristimerin	antineoplásico, antiinflamatório
quinalizarina	antiviral, inibidor da integrase do HIV-1
rifampicina	antibacteriano
rifaximina	antibacteriano, inibidor da síntese de RNA
desidrocolato de sódio	**colerético**
estreptozocina	antineoplásico, agente alquilante
tirotricina	antibacteriano tropical

Fonte: Adaptado de FRITZLER; ZHU (2011)

3.4 Utilização de ferramentas de bioinformática aplicadas na predição de estruturas proteicas e busca de inibidores

3.4.1 Modelagem Molecular de Proteínas: Modelagem por Homologia, *Ab initio*, *Threading* (I-Tasser)

O conhecimento das estruturas tridimensionais (3D) de proteínas pode fornecer informações importantes sobre sua função no nível molecular, uma vez que, sua função química está relacionada a uma estrutura específica. Dentre suas funcionalidades, os complexos proteicos são vitais para muitos processos celulares atuando como enzimas, transportadores, mensageiros, inibidores, fatores de transcrição, entre outros; por isso a compreensão das interações e funções bioquímicas de proteínas é fundamental, pois possibilita a descoberta de prováveis alvos moleculares, e o entendimento de processos naturais e patológicos (GORDON; KHANA, 2010; SANTOS FILHO; BICCA DE ALENCASTRO, 2003; RANGEL, 2019).

A predição de estruturas proteicas pode ser determinada utilizando a Modelagem Molecular (MM) que é considerada uma técnica computacional útil na construção de modelos tridimensionais a partir de uma estrutura molde (resolvida por cristalografia de raio-X, NMR, por exemplo), que tem contribuído para elucidar interações intra/inter moleculares, mecanismos de reações químicas, como também, a estrutura e função de proteínas difíceis de serem purificadas em larga escala (GORDON; KHANA, 2010; SANTOS FILHO; BICCA DE ALENCASTRO, 2003; RANGEL, 2019)

Baseado em princípios básicos da Mecânica Quântica e na Mecânica Molecular, com esse método, é possível realizar a construção do modelo tridimensional de uma proteína utilizando dos padrões de distância geométrica de átomos ou grupo de átomos e de ligações químicas (tamanho e ângulos), informações sobre a geometria de peptídeos (corrigidos posteriormente por algum método de otimização do modelo construído) e dados obtidos de experimentos físicos, geralmente cristalografia com difração de raios-X, difração de nêutrons ou estruturas deduzidas de ressonância nuclear magnética (NMR) (KOLLMAN, 1987; GORDON; KHANA, 2010; SANTOS FILHO; BICCA DE ALENCASTRO, 2003; RANGEL, 2019). Existem basicamente três técnicas de modelagem molecular de proteínas, modelagem comparativa, *ab initio* e *Threading*.

A técnica de Modelagem Comparativa ou também chamada de Modelagem por Homologia, se baseia no alinhamento dos aminoácidos idênticos entre a proteína-alvo e a proteína-molde e no compartilhamento da conformação espacial entre esses resíduos. Portanto, a técnica computacional que apresenta maior acurácia na determinação *in sílico* de estruturas proteicas é a técnica de modelagem comparativa ou por homologia, porém sua principal desvantagem está relacionada a limitação da dependência de uma estrutura-molde que seja

previamente determinada (GORDON; KHANA, 2010; SANTOS FILHO; BICCA DE ALENCASTRO, 2003).

Entretanto, a Modelagem Comparativa (por Homologia) de proteínas fundamenta-se no fato de que as sequências evolutivamente relacionadas possam compartilhar o mesmo padrão de enovelamento da estrutura terciária, ou seja, possui o princípio da conservação de estruturas entre proteínas homólogas. Com base nesta premissa, geralmente, o processo de obtenção de um modelo proteico virtual através da execução da estratégia que envolve cinco etapas principais, esquematizados na figura 11: 1) identificação, alinhamento e seleção do molde; 2) construção do modelo; 3) refinamento; 4) modelagem das alças; 5) modelagem das cadeias laterais; e 6) validação (CALIXTO et al., 2013).

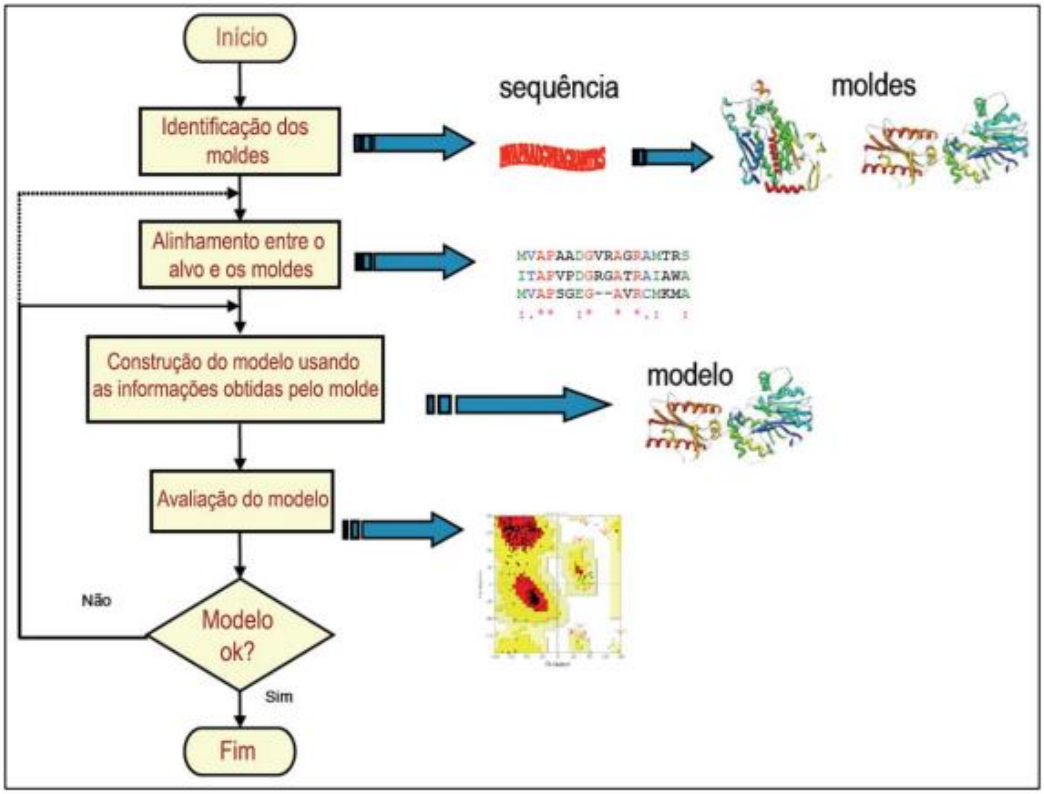

Figura 11: Esquema geral do método teórico da modelagem comparativa. Fonte: Adaptado de MARTÝ-RENOM et al., (2000).

Portanto, a acurácia da estrutura gerada por esta técnica é diretamente proporcional ao grau de identidade que ocorre com a proteína-molde, caso esse percentual de identidade seja maior ou igual a 50%, o modelo gerado apresentará de boa à ótima qualidade e com resultados precisos, se o percentual oscila entre 30 e 50% o modelo obtido apresenta boa qualidade, e percentual de identidade inferior a 30% podem ocorrer devido a erros no alinhamento, a estrutura gerada apresentará baixa qualidade e difícil correção no estágio de refinamento do

modelo, que consiste principalmente em ajustar a molécula, para que assuma menor estado de energia livre (minimização de energia), utilizando programas de dinâmica molecular como CHARM, AMBER OU GROMOS. Em casos, do grau de identidade inferior a 25%, faz-se necessário encontrar parâmetros experimentais e modificar o alinhamento (KOPP; SCHWEDE, 2004; FRENKEL; TRIFONOV, 2007; SANTOS FILHO; BICCA DE ALENCASTRO, 2003)

O SWISS-MODEL Workspace (http://swissmodel.expasy.org) (WATERHOUSE et al., 2018) é um servidor automatizado e gratuito para predição de estruturas de proteínas via Modelagem por Homologia, os alinhamentos múltiplos do molde são utilizados afim de identificar regiões estruturais conservadas que compartilham a mesma família ou não, e que tenham semelhanças de função e dobra, que são modeladas com o método de corpos rígidos, enquanto que as regiões variáveis são construídas em torno deles. A construção de um modelo ocorre baseando-se no fato de que proteínas homólogas possuem regiões estruturalmente conservadas como α-hélices e folhas β. Utiliza como método de avaliação e validação das estruturas modelos construídas as estimativas de qualidade pelo QMEAN e pelo Z-escore (SCHWEDE et al., 2003; MANDHAVI SASTRY et al., 2013; WATERHOUSE et al., 2018).

Uma outra técnica de modelagem de proteínas é a técnica *ab initio* (de novo) que se baseia em princípios físicos de energia, ao invés do uso de moldes (ou seja, sem a influência de proteínas homólogas); contudo, só é possível aplicá-la a proteínas com no máximo 200 aminoácidos. Há métodos que tentam mimetizar o processo de enovelamento de proteínas e outros que aplicam métodos estocásticos para pesquisar possíveis soluções estruturais. Sua desvantagem é devido a utilização de métodos computacionalmente muito exigentes, porém foram aplicados com sucesso para determinação de proteínas pequenas. O servidor Web ROBETTA (http://robetta.bakerlab.org) utiliza a implementação automatizada do programa ROSETTA (https://www.rosettacommons.org/) no qual é possível realizar tanto modelagem comparativa quanto *ab initio* (SONG et al., 2013)

O Servidor I-Tasser (ROY; KUCUKURAL; ZHANG, 2010) também utilizado para realizar modelagem molecular é um software que utiliza da metodologia do método de modelagem por homologia e do método *ab initio* denominado *threading*, que tem o objetivo de modelar estruturas que possuem enovelamento similar a proteínas de estruturas conhecidas (através do reconhecimento de padrões de enovelamento), e que compartilham baixo grau de similaridade. Neste método, é bastante empregado quando possuem modelos de baixa identidade que cobrem regiões diferentes da sequência alvo, uma vez que, a sequência é fragmentada na busca por homólogos estruturais, explorando muitos

alinhamentos; uma das vantagens mais importantes do I-Tasser é que nessa montagem de fragmentos pode conduzir as estruturas de modelo iniciais para mais perto de seus estados nativos e podendo gerar modelos precisos de proteínas (VERLI, 2014; ZHANG, 2009; WU, S.; ZHANG, 2007)

Portanto, o procedimento para gerar modelos precisos de proteínas utilizando o I-Tasser (*threading*) baseia-se em realizar o alinhamento de múltiplos fragmentos com a estrutura tridimensional conhecida depositada em banco de dados de estruturas de proteínas, com a sequência de interesse, considerando as semelhanças da estrutura secundária predita para a proteína de interesse e a do fragmento molde na mesma região. Em seguida, os fragmentos gerados são agrupados, e a técnica *ab initio* é aplicada para as regiões não alinhadas; os fragmentos são organizados em estrutura 3D através de simulações de Monte Calor (troca entre réplicas) até que construa um modelo otimizado. Por fim, é realizado o refinamento em que são recuperados os principais confôrmeros obtidos no passo anterior, para realizar mais uma vez o agrupamento por troca de réplicas afim de eliminar impedimentos estéricos, resultando no final desta etapa em cinco modelos que serão adicionadas cadeias laterais e submetido ao último passo de refinamento para otimização dos rotâmero (ROY; KUCUKURAL; ZHANG, 2010; WU, S.; ZHANG, 2007; ZHANG, 2009)

Para avaliar e validação da precisão da estrutura predita, uma pontuação de confiança chamada de C-*score* é definida com base na qualidade dos alinhamentos de enovelamento (*threading*) e na convergência das simulações de refinamento de montagem estrutural do I-TASSER e para avaliar a função da estrutura construída, a pontuação de confiança é definida pelo C-*score* da previsão de estrutura e na similaridade estrutural global e/ou local entre os modelos previstos e seus análogos estruturais no PDB (ROY; KUCUKURAL; ZHANG, 2010)

3.4.2 Triagem virtual baseada na estrutura do ligante (LBVS) e Triagem virtual baseada na estrutura do receptor (RBVS)

A Triagem Virtual consiste em selecionar as melhores moléculas a partir de uma grande variedade de técnicas computacionais (*in sílico*), possui grande vantagem em relação as análises *in vitro* e *in vivo*, devido sua capacidade de realizar análises mais rápidas e menos dispendiosas. Dependendo dos dados disponibilizados, pode-se realizar uma triagem virtual baseada no alvo (ou receptor), caso este seja conhecido, ou ainda, baseado no ligante, analisando os dados de várias moléculas e com a utilização de técnicas de comparação entres

elas, sem a necessidade de usar os dados do alvo. Como também é possível utilizar ambas as técnicas em conjunto, que analisam tanto o alvo quanto os possíveis ligantes, resultando numa abordagem mista, porém é necessário o conhecimento da molécula-alvo e de alguns ligantes dessa molécula para esta abordagem (LINHARES et al., 2021; RANGEL, 2019; DOMINGUES, 2012; VELOSO, 2019)

Na triagem virtual baseada no receptor (*Receptor-Based Virtual Screening* – RBVS) ou também conhecida como triagem virtual baseada em alvos (*Target-Based Virtual Screening* – TBVS), esquematizada na figura 12, fundamenta-se no conhecimento da estrutura da proteína-receptor (alvo) para que ocorra a busca na descoberta de ligantes prováveis. Dessa maneira, os algoritmos que selecionam os ligantes, dentre um conjunto de uma variedade de moléculas, dependem das informações acerca do receptor para estimar as probabilidades de interação (DOMINGUES, 2012; CORTES-CABRERA et al., 2016; LINHARES et al., 2021; RANGEL, 2019; VELOSO, 2019)

De modo diferente, a triagem baseada em ligantes (*Ligand-Based Virtual Screening* – LBVS) (figura 12) não depende do receptor (alvo), uma vez que, fundamenta-se em buscar novas moléculas candidatas a ligantes (focalizando nas similaridades estruturais e nas atividades moleculares) a partir de ligantes já conhecidos e disponíveis em bancos de dados virtuais; portanto, os descritores que envolvem os próprios ligantes são utilizados como discriminantes nos algoritmos de seleção e também de classificação, baseando-se em informações de um conjuntos de ligantes ativos e também não ativos para uma determinada função, porém esta técnica é promissora ou única opção quando o receptor (alvo) não é bem caracterizado ou conhecido (DOMINGUES, 2012; CORTES-CABRERA et al., 2016; LINHARES et al., 2021; RANGEL, 2019; VELOSO, 2019)

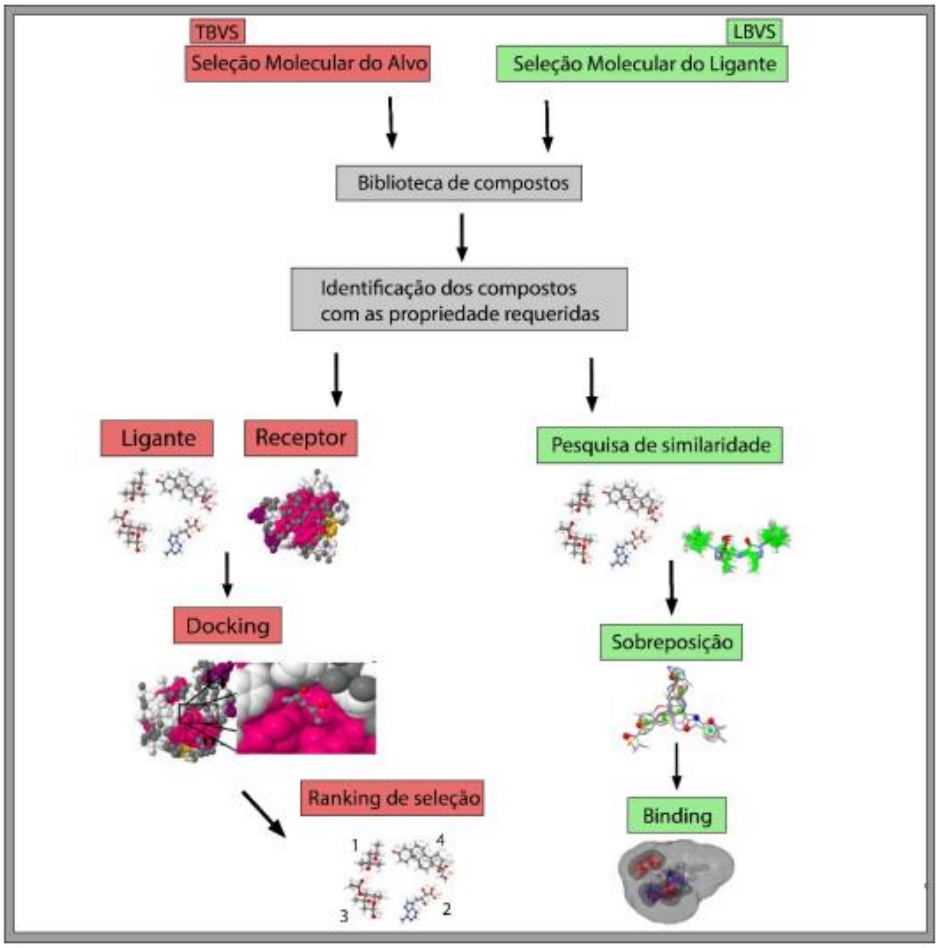

Figura 12: Esquematização das abordagens TBVS e LBVS. Fonte: LINHARES et al., (2021).

Para a finalidade deste trabalho, utilizamos a técnica de triagem virtual baseada no receptor (RBVS) , devido a presença de informação estrutural e experimental sobre o alvo de interesse. O conhecimento do sítio de ligação (*binding site ou pocket*) foi útil para a construção de um ligante no próprio sítio da proteína alvo, e realização de estudos do sítio alvo para estimarmos computacionalmente, a afinidade com outros ligantes prováveis.

3.4.3 Acoplamento Molecular (*Molecular Docking*)

A técnica de acoplamento molecular, também chamada de *Docking* molecular, possibilita complementar as técnicas de triagem virtual, fornecendo novos parâmetros para a afinidade dos ligantes prováveis com a molécula-alvo.

Segundo Morris e Lim-Wilby, 2008 o objetivo principal do software de acoplamento molecular (*docking*) automatizado é “entender e prever o reconhecimento molecular, tanto estruturalmente, encontrando modos de ligação prováveis, quanto energeticamente, prevendo a

afinidade de ligação” , os principais constituintes do acoplamento molecular (*docking*) incluem (figura 13): a estrutura tridimensional (3D) de uma molécula-alvo (ou proteína) e a constituição molecular de uma pequena molécula (ligante), que irão interagir; a definição do sítio ativo (*binding site* ou *pocket*) alvo; um algoritmo de busca para gerar *poses* de ligação significantes do ligante ao sítio ativo alvo que foi definido; e um método de pontuação (*score*) útil para avaliar as configurações geradas envolvendo sua posição, orientação, e conformação do complexo formado entre proteína-ligante, sendo este acoplamento molecular ideal quando utilizado após uma prévia seleção das moléculas a serem envolvidas (triagem virtual) (WARTERS; STAHL; MURCKO, 1998; TROTT & OLSON, 2010)

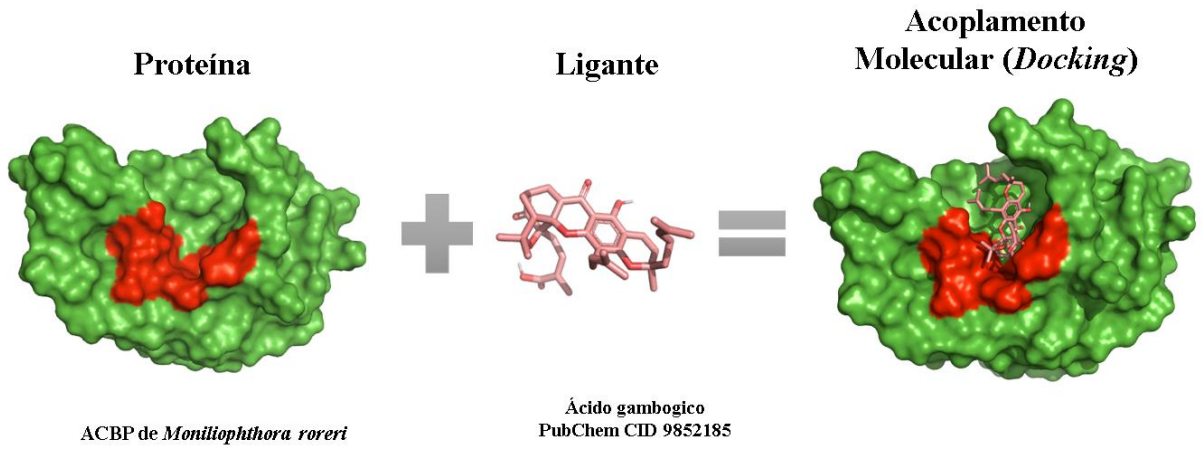

Figura 13: Representação esquemática da formação do acoplamento molecular (*docking*) entre uma proteína-alvo interagindo com o ligante e gerando um complexo, com a região do sítio ativo destacado em vermelho. Fonte: Imagem gerada pelo PyMol, SCHRÖDINGER (2021)

Em princípio, a preparação para o acoplamento molecular (*docking*) se dar com a obtenção da estrutura tridimensional (3D) da molécula-alvo e/ou ligante por técnicas como ressonância magnética nuclear (*Nuclear Magnetic Ressonance* – NMR), cristalografia por difração de raios-X ou de nêutrons, a partir dos bancos de dados de acesso público como o RCSB PDB – *Protein Data Bank* (BERMAN et al., 2000) ou ainda a partir da criação/modelagem da estrutura da molécula-alvo de interesse, devendo-se atentar para a qualidade da mesma. Quando se tratando da cristalografia, deve-se levar em consideração a resolução e outros parâmetros que aferem a qualidade dos dados gerados (WARTERS; STAHL; MURCKO, 1998; WARREN et al., 2006; TROTT; OLSON, 2010; VELOSO, 2019).

A realização do acoplamento molecular (*docking*) pode ocorrer por uma área de busca do algoritmo que esteja limitada à uma região específica, de preferência no sítio ativo. Além

disso, muitas das estruturas de proteínas armazenadas no *Protein Data Bank* (PDB) constituem de moléculas de água em sua superfície que podem consequentemente, ou não obstruir o sítio ativo e, logo, deve-se considerar a remoção destas moléculas. Outra consideração importante, se deve a cristalografia por difração de raios-X, que geralmente, não torna possível observar átomos de hidrogênio, porém, é possível que sejam adicionados por programas específicos, que admitem também o pH e contexto químico-estrutural do alvo a ser protonado. Frequentemente, os possíveis ligantes que serão obtidos de uma base de dados com somente a representação atômica em suas dimensões, serão necessários realizar a conversão para o modelo de estrutura tridimensional (3D), com uso de programas, para visualizar suas diversas conformações possíveis ou mais prováveis (WARTERS; STAHL; MURCKO, 1998; WARREN et al., 2006; TROTT; OLSON, 2010; VELOSO, 2019).

Contudo, uma vez com as estruturas individuais, o objetivo é que possam encontrar uma combinação de complexos alvo-ligantes que possuam a melhor pontuação (*score*); portanto, os métodos de avaliação da energia livre de ligação têm a finalidade de otimizar a posição, orientação e conformação do ligante baseados numa função objetiva, e que levam em consideração tanto as interações físico-químicas quanto, adequações geométricas entre as moléculas. Entre as principais interações intermoleculares envolvidas nos complexos alvo-ligante estão: ligações de hidrogênio, iônicas, interações hidrofóbicas (ou forças de van der Waals), empilhamentos aromáticos, cátion—Pi e coordenação com íons metálicos (MORRIS; LIM-WILBY, 2008; RANGEL, 2019; WARREN et al., 2006; TROTT; OLSON, 2010).

O pacote MGL Tools que contém os programas AutoDock Tools (MORRIS et al., 2009) e AutoDock Vina (TROTT; OLSON, 2010) que consiste nos dois principais softwares mais utilizados para realizar o acoplamento molecular (*docking*), que possibilitam estimar uma *pose* e energia de ligação entre uma proteína-alvo e uma molécula (geralmente pequena) denominada ligante; esses programas têm contribuído para a descoberta de vários fármacos e inibidores (MORRIS; LIM-WILBY, 2008; WARREN et al., 2006)

O programa AutoDock Vina (TROTT; OLSON, 2010), possui praticidade, rapidez na geração dos resultados e precisão do modo de ligação entre as moléculas envolvidas, possibilitando para cada posição (*pose*) gerar uma pontuação (*score*), que leva em consideração um campo de força de energia livre semi-empírico (figura 14) segundo Huey et al., 2007, cujo objetivo é capturar as contribuições entálpicas e entrópicas (físico-químicas) do sistema utilizando de termos aos pares para estimar a energia da ligação de duas ou ainda mais moléculas situadas no ambiente aquático e desse modo avaliar a interação entre as duas

moléculas, além de estimar a contribuição da água circundante através de um método empírico. Diferentemente do campo de força da mecânica molecular tradicional, que usa moléculas de água para avaliar as contribuições de solvatação.

Consequentemente, estima-se que (figura 14) a energia livre de ligação ocorra como resultado da diferença entre a energia do ligante e a proteína em um estado não ligado separados e a energia do complexo formado entre o proteína-ligante (estado ligado), possibilitando a avaliação da energia intramolecular da transição do estado não ligado para a conformação ligada para cada uma das moléculas separadamente e, em seguida, avalia a energia intermolecular da junção das duas moléculas no complexo ligado. Neste exemplo, supõe-se que as duas moléculas estão suficientemente distantes uma da outra no estado não ligado para que a energia do complexo não ligado seja zero, e que não permitiu movimento na proteína, então o estado ligado da proteína é idêntico ao estado não ligado da proteína, e a diferença em sua energia intramolecular é zero (HUEY et al., 2007).

$$\Delta G = (V_{\text{bound}}^{\text{L-L}} - V_{\text{unbound}}^{\text{L-L}}) + (V_{\text{bound}}^{\text{P-P}} - V_{\text{unbound}}^{\text{P-P}}) + (V_{\text{bound}}^{\text{P-L}} - V_{\text{unbound}}^{\text{P-L}} + \Delta S_{\text{conf}})$$

Figura 14: Equações da abordagem no AutoDock. Representando o Campo de força de energia livre semi-empírico com contribuições entálpicas e entrópicas (fisíco-químicas) do sistema utilizando de avaliações dos termos aos pares para estimar a energia da ligação. **Abreviaturas:** ΔG:estimativa da energia livre de ligação, V^{L-L}bound: energia livre ligado, V^{L-L} unbound: energia do ligante não ligado, V^{P-P} bound: energia da proteína ligada, V^{P-P} unbound: energia da proteína não ligada, V^{P-L} bound: energia do complexo ligado, V^{P-L} unbound: energia do complexo não ligado, ΔS_{conf}: estimativa da entropia conformacional. **Fonte:** HUEYet al., (2007).

Na figura 15, representa a equação referente a função de pontuação do AutoDock Vina (TROTT; OLSON, 2010), que leva em consideração as interações (ou forças) de van der Waals, ligações de hidrogênio, interações estéricas (gauss 1, gauss 2, e repulsão), hidrofobicidade, e número de ângulos de torção de rotação livre dos ligantes (TROTT; OLSON, 2010; HUEYet al., 2007).

$$V = W_{vdw} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) + W_{\text{hbound}} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{\text{elec}} \sum_{i,j} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} + W_{\text{sol}} \sum_{i,j} (S_i V_j + S_j V_i) e^{(-r_{ij}^2/2\sigma^2)}$$

Figura 15: Equação da função de pontuação do AutoDock Vina. Abreviaturas: W: constante de ponderação, vdw: Van der Waals, hbound:ligação de hidrogênio, elec: eletroestática, sol: parâmetro de solvatação. A e B são parâmetros do campo de força AMBER, C e D são parâmetros para atribuir profundidade focado em ligações de hidrogênio. E(*t*):direcionalidade da ligação de hidrogênio, qiqj:potencial de Coulomb. V: volume. **Fonte:** HUEYet al., (2007).

Portanto, o campo de força de energia livre semi-empírico com contribuições entálpicas e entrópicas em conjunto com um algoritmo genético lamarckiano, fornecem a previsão de conformações ligadas com energias livres de associação previstas, que podem ser calculados pela função de pontuação a partir de potenciais que envolvem as ligações de van der Waals e eletrostáticas, e outros termos que irão assumir sistemas de massa-mola em energias de ligação (estimadas com valores de aproximação), as angulares e desvios do planos. Seu algoritmo de optimização global é o método de *Iterated Local Search* (Busca Local Iterada), que realiza a busca das melhores *poses*, que implica em menor energia livre, ou seja, esta metodologia tem por finalidade encontrar o modo de ligação com a menor energia de interação possível entre proteína-ligante (TROTT; OLSON, 2010; MORRIS; LIM-WILBY, 2008; WARREN et al., 2006; RANGEL, 2019).

A parte da superfície que este programa realiza a busca é especificada a partir das coordenadas da "caixa" (*grid box*), preparado na função do programa AutoDock Tools (MORRIS et al., 2009) (figura 16), essas coordenadas irão determinar o espaço der busca do algoritmo. Com a busca iniciada a partir de um valor numérico aleatório ou conforme disponibilizado pelo usuário, possibilita a reprodução do experimento. Os arquivos de entrada para este programa incluem o receptor e o ligante nos formatos PDBQT, além de necessitar das coordenadas *box* (que equivale ao tamanho e posição no espaço tridimensional), número máximo de *poses* a serem geradas (geralmente, são 9 *poses*), variação de energia (*energy range)* é o valor máximo de energia entre a melhor pose que foi predita e a pior, dentre o número máximo de *poses*, e *exhaustiveness* (quantidade de repetições e os números de cálculos simultâneos = valor de *exhaustiveness*). Em suma, as etapas principais do docking incluem: 1) seleção e preparação do alvo, 2) seleção e preparação do ligante, 3) configuração do *docking* e

4) avaliação dos resultados do *docking* molecular (MORRIS; LIM-WILBY, 2008; TROTT; OLSON, 2010; WARREN et al., 2006).

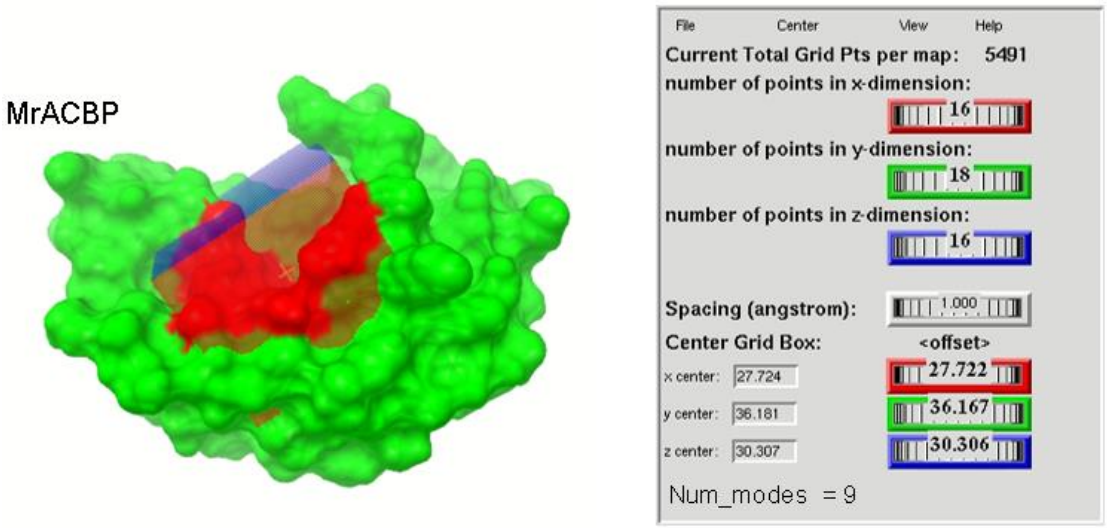

Figura 16: Construção da "caixa" (*grid box*) com coordenadas de x, y e z do sítio ativo do receptor MrACBP para realizar o acoplamento molecular (*docking*). Na região central da proteína mostra a caixa (*grid box*) no meio do bolso de ligação. Ao lado direito da imagem as configurações das coordenadas de x, y e z para ser utilizado no AutoDock Vina, como local de interação com ligantes. Num_modes: se refere a quantidade de graus de liberdade, neste caso 9 *poses*, para as ligações intermoleculares. **Fonte:** Imagem gerada no AutoDock Tools v.1.5.6; MORRIS et al., (2009).

3.4.4 Simulações de Dinâmica Molecular (DM): Aspectos gerais e aplicações

A Dinâmica Molecular é uma técnica responsável por simular computacionalmente o comportamento molecular por um período de tempo. Logo, as simulações computacionais dos sistemas biológicos têm se tornado uma ferramenta muito útil nas últimas décadas, possibilitando obter dados teóricos da movimentação estrutural e também, das interações intra e intermoleculares de biomoléculas antes de serem avaliados por métodos experimentais (DURRANT; McCAMMOM, 2011; HOSPITAL et al., 2015; KARPLUS; McCAMMON, 2002).

Para o estudo da relação de interação proteína-ligante, a metodologia aplicada nas simulações de DM, é embasado na utilização da Mecânica Molecular (MM, dinâmica clássica ou mecânica clássica), que se baseia na integração das leis de movimento de Newton (HOLLINGSWORTH; DROR, 2018). Essa metodologia tem a vantagem de possuir simplicidade computacional, eficiência, e tem sido o mais utilizado para que se possa modelar estruturas moleculares e energias conformacionais para elucidação de problemas experimentais. Sua metodologia se baseia na aproximação de Born-Oppenheimer, dentro dessa

aproximação considera de forma separadas os movimentos nucleares e dos elétrons; uma vez que, na mecânica molecular supõe que os elétrons possam se ajustar de forma instantânea a qualquer movimentação livre que ocorre nos núcleos, sendo essas interações aditivas em relação ao núcleo. A superfície de Born-Oppenheimer ou superfície de energia potencial, é referida como a energia de uma molécula em termos das posições assumidas pelos movimentos nucleares (multidimensional), além disso, nessa superfície contém mínimos locais que ocorrem aproximação de interações em relação a uma ou algumas conformações moleculares estáveis, e numa primeira aproximação, a molécula é descrita pela estrutura que a corresponde ao mínimo global da superfície de energia potencial, que será calculada usando o campo de força (COELHO et al., 1999)

As equações do Campo de força denotam a energia potencial e as forças de características moleculares. Na Mecânica Molecular considera as moléculas constituídas de átomos ligados entre si por forças elásticas ou harmônicas. Na figura 17, mostra as forças atômicas que contribuem com o movimento molecular, e que surgem das interações entre (A) átomos quimicamente ligados, e (B) não ligados, são elas: alongamento das ligações atômicas (modeladas com molas virtuais de comprimento), os ângulos e rotações em torno de uma ligação (diedros) que são modelados através da aproximação das diferenças de energia entre conformações. As Forças não ligadas surgem devido a interações de van der Waals, e interações carregadas (eletrostáticas), para modelar estas forças são utilizados o potencial de Lennard-Jones e lei de Coulomb, respectivamente (DURRANT; McCAMMON, 2011; HOSPITAL et al., 2015; COELHO et al., 1999)

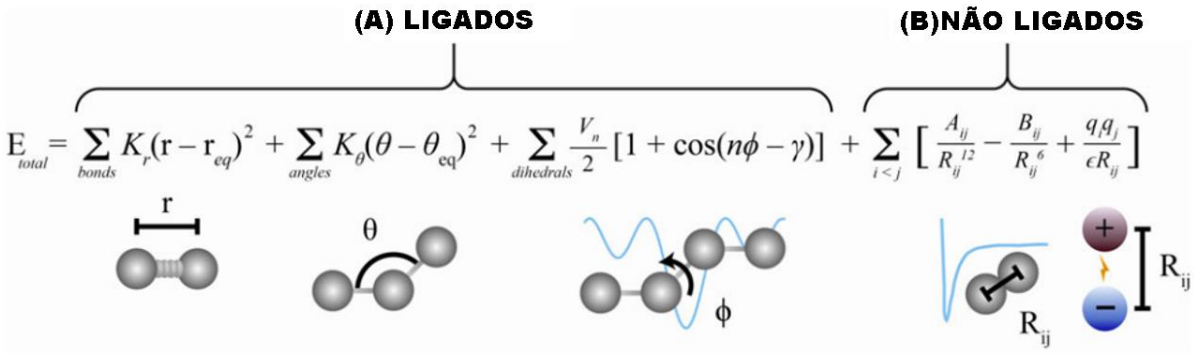

Figura 17: Equação para aproximar as forças atômicas que atuam em cada átomo e que governam o movimento molecular. Surgem das interações entre: (A) átomos quimicamente ligados e, (B) Não ligados; são elas: alongamento das ligações atômicas, os ângulos e rotações em torno de uma ligação (diedros). As Forças não ligadas surgem devido a interações de van der Waals, e interações carregadas (eletrostáticas) **Fonte:** DURRANT; McCAMMON (2011).

A esquematização da abordagem das simulações de DM (figura 18) envolvem os seguintes passos: (1) Inicialmente, utilizando uma ferramenta computacional é preparado um modelo do ligante-receptor, (2) calcular repetidamente as forças através das equações que se referem ao Campo de Força de mecânica molecular, mostrada na figura (B) é usada para estimar as forças que atuam em cada um dos átomos do sistema, pela qual é feita a predição da energia potencial dos átomos que constituem a estrutura molecular. Em seguida, (3) usando essas forças para atualizar a posição dos átomos que se deslocam de acordo com as leis do movimento de Newton e a velocidade de cada átomo resultando numa trajetória; afim de, (4) prever como cada átomo em uma proteína ou outro sistema molecular se movimenta ao longo do tempo (os passos de tempo são geralmente de alguns femtossegundos (1 ou 2 fs) cada, para manter estabilidade numérica) e o processo ocorre repetidamente (DURRANT; McCAMMON, 2011; KARPLUS; McCAMMON, 2002; HOLLINGSWORTH; DROR, 2018; DU et al., 2016; HOSPITAL et al., 2015)

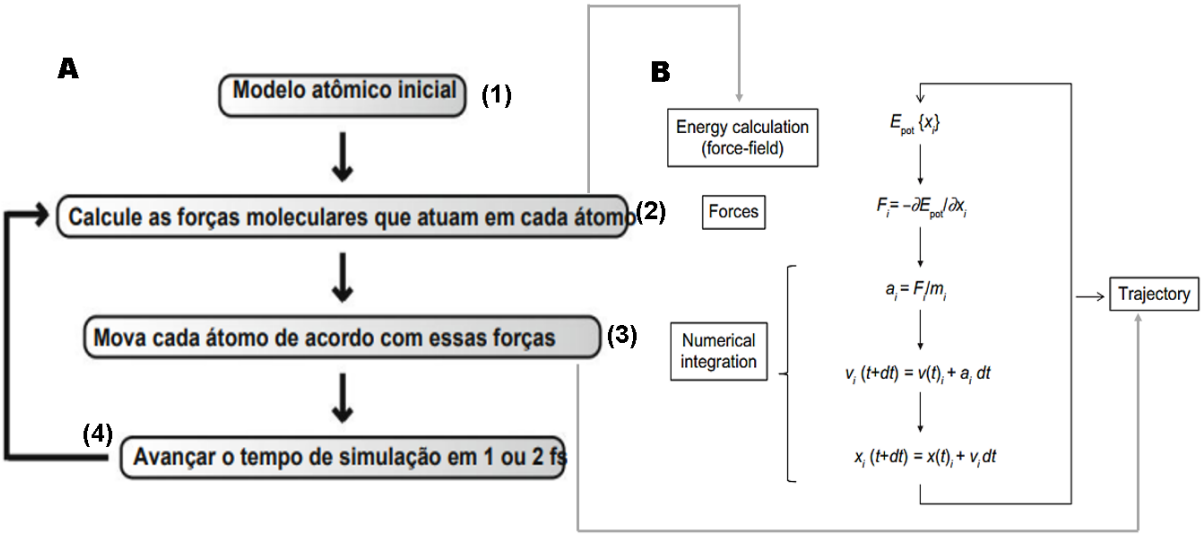

Figura 18: A. Esquematização da abordagem da simulação de Dinâmica Molecular. B. Algoritmo básico de DM. (1) Modelo atômico inicial, (2) Cálculo das forças moleculares que atuam em cada átomos, (3) Movimento de cada átomo de acordo com as forças intermoleculares, (4) O tempo de simulação é avançado (1 ou 2 fs) e o processo é repetido várias vezes. Abreviaturas (B) Epot: energia potencial, x: coordenada do átomo, F: forças, ∂Epot: derivada da energia potencial, ∂xi: derivada da posição inicial, a: aceleração, m: massa do átomo, v: velocidade, t: tempo, dt: derivada do tempo, x(t): posição em função do tempo de simulação. **Fonte:** Adaptado de DURRANT; McCAMMON (2011); HOSPITAL et al. (2015)

Ao longo das últimas décadas, ocorreram avanços contínuos na arquitetura de computadores e em softwares para simulações de DM, essas melhorias foram causadas em grande parte pela utilização de computadores de alto desempenho (HPC) e da simplicidade do algoritmo DM básico (Figura 18 B) que permitiram melhorias significativas no desempenho das simulações de DM, em virtude disso, estimulou o extenso crescimento do seu uso como

ferramenta principal nos estudos computacionais de proteínas, bem como uma maquinaria complementar em estudos experimentais, impulsionando ações para melhorar vários aspectos como velocidade, precisão e acessibilidade da simulação, e garantindo aumento da qualidade do poder preditivo e análise *in silico* mais confiável de estruturas de proteínas, dinâmica e funções (HOSPITAL et al, 2015; LAZIM; SUH; CHOI, 2020; HOLLINGSWORTH; DROR, 2018)

Atualmente, os campos de força modernos usados em simulações moleculares atomísticas são equivalentes, entre os mais utilizados incluem o AMBER, GROMOS, e o CHARMM (HUANG et al., 2016); utilizados para calcular acelerações e velocidades e para atualizar as posições dos átomos, porém diferem na forma como são parametrizados e nem todos os campos de força permitem representar todos os tipos de moléculas (HOSPITAL et al., 2015).

Quando usada de forma eficiente no estudo da dinâmica de proteínas no nível atômico se torna uma ferramenta teórica rigorosa que fornece respostas confiáveis em relação a estrutura-função de proteínas (LAZIM; SUH; CHOI, 2020) Em suma, a metodologia das simulações de DM, se baseia na aplicação das leis do movimento de Newton para prever a posição espacial de cada átomo em função do tempo (HOLLINGSWORTH; DROR, 2018) Com base nessa premissa, muitos estudos de simulação de DM visam observar uma ampla variedade de processos biomoleculares importantes ao longo de um período de tempo (figura 19), incluindo,(a) transporte através de uma membrana – translocação proteica, (b) alteração conformacional induzida pela ação de uma molécula ligante (alosteria) que pode ser um ativador, inibidor ou ambos ou ainda provocada por voltagem, (c) dobramento de proteínas e (d) ligação de ligantes, revelando as suas posições e o movimento em relação ao sítio de ligação e interação com receptor temporal (HOLLINGSWORTH; DROR, 2018; LAZIM; SUH; CHOI, 2020)

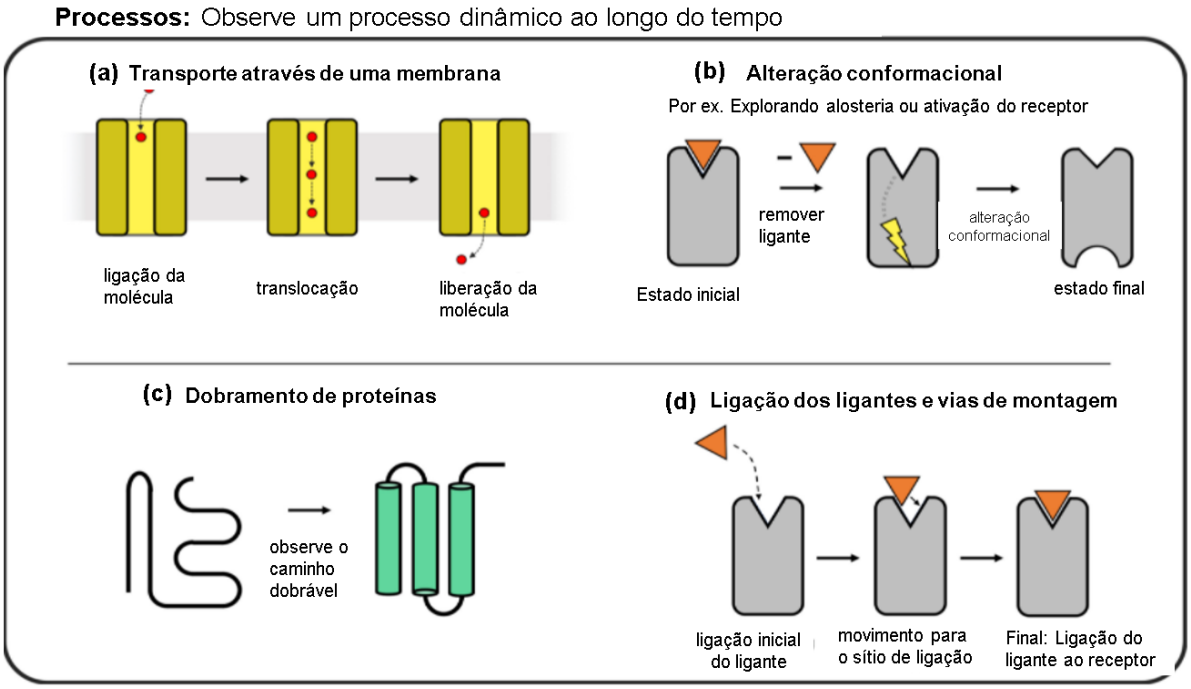

Figura 19: Ilustração da variedade de processos biomoleculares dinâmico ao longo do tempo que podem ser observados a partir das simulações de DM. (a) Transporte de membrana, (b) alteração conformacional, (c) dobramento de proteínas e (d) ligação dos ligantes e vias de montagem. **Fonte:** Adaptado de HOLLINGSWORTH; DROR (2018).

As ferramentas de simulação de DM possibilitam gerar dados da dinâmica de proteínas que podem ser traduzidos em estatísticas úteis para a previsão, caracterização e elucidação precisas dos mecanismos responsáveis pela modulação de processos biológicos, como ligação proteína-ligante e da associação proteína-proteína, que podem ser exploradas a partir da termodinâmica e cinética. Um objetivo final das simulações de DM consiste em reproduzir os mesmos resultados termodinâmicos e cinéticos, que seriam encontrados em experimentos, mas com um número maior de detalhes atomísticos. Visto que, ao comparar simulações realizadas em diferentes condições de temperatura, pH, pressão, portanto, pode-se identificar os efeitos de uma ampla variedade de perturbações moleculares. A compreensão de uma estrutura de nível atômico das biomoléculas possibilita gerar uma visão substancial sobre como elas funcionam, ou seja, prever como as biomoléculas responderão em nível atômico, a perturbações como mutação, fosforilação, protonação ou adição ou remoção de um ligante (LAZIM; SUH; CHOI, 2020; HOLLINGSWORTH; DROR, 2018)

Dentre as aplicações das simulações de DM, ilustrada na figura 20, a mais comum e intuitiva consiste em avaliar a estrutura e dinâmica de uma biomolécula, estudando sua flexibilidade conformacional e estabilidade, permitindo quantificar quais os tipos de flutuações estruturais sofrem e o quanto várias regiões da molécula se movem em equilíbrio. Podendo

ainda revelar o comportamento dinâmico de moléculas de água e íons de sal, que são frequentemente críticos para a função de proteínas e ligação de ligantes (LI et al., 2013).

Outra aplicação das simulações de DM (figura 20) é testar a precisão de uma estrutura modelada ou mesmo para o refinamento de modelos estruturais oriundos da modelagem por homologia de proteínas, de um modelo inicial com alta energia para um modelo de menor energia das interações intermoleculares, como também, através de uma simulação da estrutura cristalina corrigir artefatos, permitindo que a estrutura alterne para uma conformação mais favorável quando esta encontra-se num ambiente apropriado de solvatação. Desse mesmo modo, também é possível testar *poses* (posições) de ligação de modelos proteína-ligante; uma *pose* que se apresente estável na simulação DM é mais provável ser considerada precisa do que uma que é instável (CLARK et al., 2016)

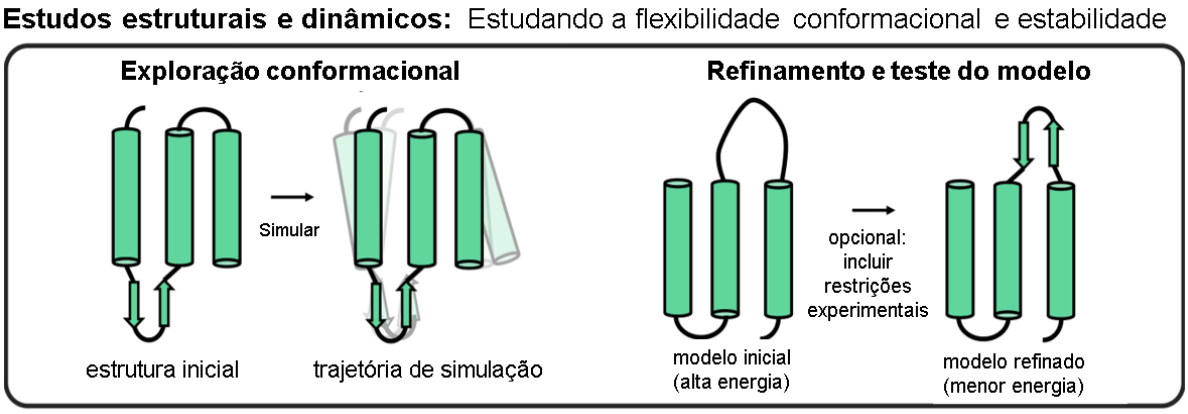

Figura 20: Aplicações de Simulações de Dinâmica Molecular nos estudos estruturais e dinâmicos de uma biomolécula, avaliando sua flexibilidade conformacional e estabilidade. Fonte: Adaptado de HOLLINGSWORTH; DROR (2018)

Por outro lado, determinar como um sistema biomolecular responderá após alguma perturbação ou mudança controlada (figura 21), é umas das mais importantes aplicações da simulação DM. Por exemplo: (a) remover um ligante ligado de uma estrutura de proteína, ou substituir um ligante que estava ligado por um ligante diferente ou ainda, adicionar um ligante na estrutura experimental que não possuía um ligante ligado, e em todos esses casos, simular para ver como afetam a conformação da proteína (PROVASI et al., 2011); (b) mutar um ou mais aminoácidos na proteína, para explicar ou prever o efeito funcional de uma mutação ou para obter a estrutura do tipo selvagem (quando modelo construído ou resolvido experimentalmente é diferente do tipo selvagem) (CORDERO-MORALES et al., 2007); (c) alterar o estado de protonação de um aminoácido (ácido ou básico) (LIU; KE; GONG, 2015); aplicar forças externas (d) deformação mecânica: puxando o ligante para fora do bolso de

ligação, ou (e) aplicar um potencial externo aos átomos simulados para capturar o efeito da voltagem transmembrana (DELEMOTTE et al., 2011); (f) alterar o ambiente molecular de uma proteína, interferindo na concentração de sal ou a composição de lipídios em uma membrana. Vale ressaltar que, faz-se necessário realizar várias simulações dos sistemas perturbado e não perturbado para que seja possível observar e identificar as diferenças nos resultados de cada um desses exemplos (HOLLINGSWORTH; DROR, 2018)

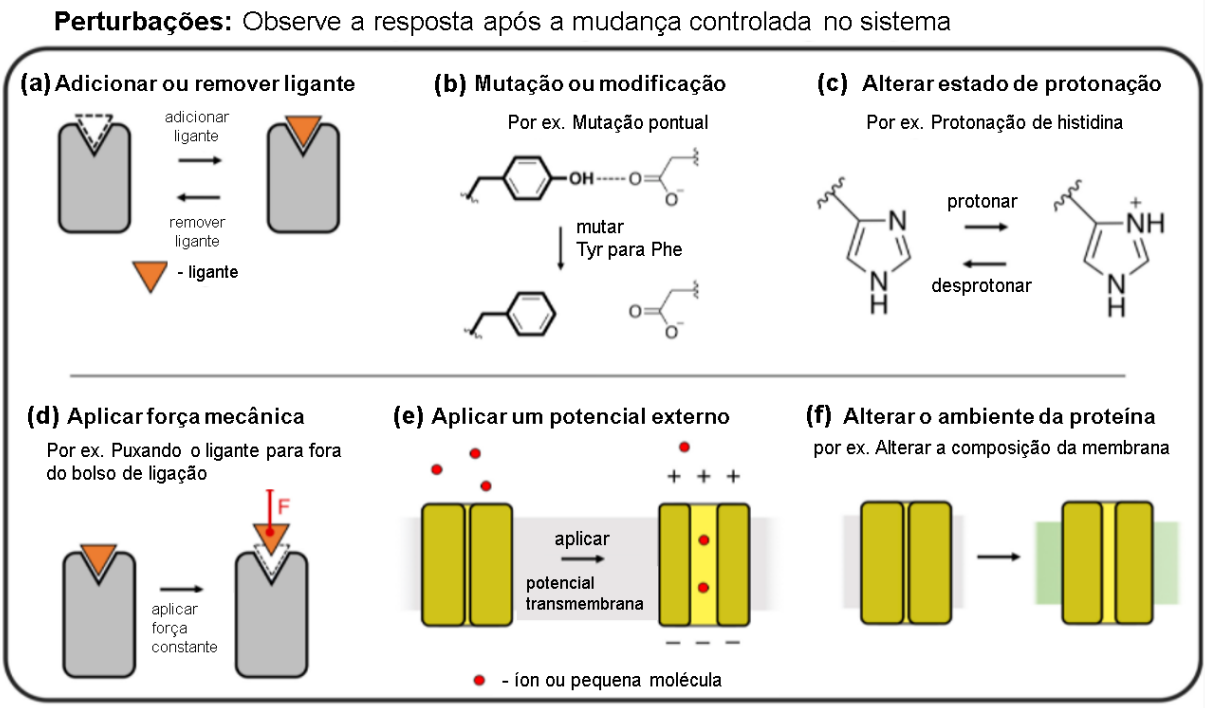

Figura 21: Aplicações de simulações de Dinâmica Molecular com alguma perturbação ou mudança controlada no sistema biomolecular. Fonte: Adaptado de HOLLINGSWORTH; DROR (2018).

4. MATERIAL E MÉTODOS

4.1 Análises *in sílico* da ACBP de *Moniliophthora roreri*

A sequência fasta dos resíduos de aminoácidos da proteína de ligação acil-CoA (ACBP) de *Moniliophthora roreri* foi obtida no banco de dados do NCBI (https://www.ncbi.nlm.nih.gov/protein) (MEINHARDT et al., 2014; MONZANI et al., 2010) Em seguida, utilizando o servidor SignalP-5.0 (http://www.cbs.dtu.dk/services/SignalP/) (PETERSEN et al., 2011) foi verificado se a sequência de aminoácidos da proteína obtém a presença de peptídeo sinal. O Compute Pi/Mw Expasy foi utilizado para a realização de predições referentes ao ponto isoelétrico e peso molecular da proteína ACBP (https://web.expasy.org/compute_pi/) (GASTEIGER et al., 2005). Para analisarmos a sequência da estrutura realizamos a predição no banco de dados de domínios conservados

(CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) (MARCHLER-BAUER et al., 2014) no NCBI e no InterPro (https://www.ebi.ac.uk/interpro/about.html) (MITCHELL et al., 2019), para identificação de domínios conservados e a família na qual a proteína em estudo está inserida.

4.2 Modelagem molecular por homologia da ACBP de *Moniliophthora roreri*

A estrutura tridimensional (3D) da ACBP de *M. roreri* foi construída por meio da modelagem por homologia, inicialmente utilizamos a sequência de aminoácidos da proteína MrACBP depositada no NCBI (https://www.ncbi.nlm.nih.gov/protein) para realizarmos a busca de proteínas homólogas. As proteínas homólogas foram selecionadas através de uma busca de sequências de aminoácidos com áreas de similaridades com a sequência de entrada, utilizando-se a ferramenta de pesquisa de alinhamento local - BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (ALTSCHUL et al., 1990; ALTSCHUL et al., 1997), restrito à sequências com estruturas depositadas no *Protein Data Bank* (PDB) (https://www.rcsb.org/), considerando-se o grau de cobertura (o valor do *score* positivamente correlacionado com a qualidade do alinhamento, ou seja, quanto maior este valor, melhor o alinhamento) e identidade (acima de 30%) dos alinhamentos. Esses *scores* da qualidade do alinhamento entre os resíduos de aminoácidos foram analisados utilizando o programa de alinhamentos múltiplos de sequências de proteínas, o T-COFFEE (http://tcoffee.crg.cat/) (DI TOMMASO et al., 2011).

Na construção do modelo da MrACBP utilizamos o servidor SWISS-MODEL Workspace (http://swissmodel.expasy.org), que realiza todo o processo de forma automatizada, gerando os moldes resultantes da sequência submetida, apresenta várias estatísticas, e ainda permite a escolha do melhor molde, além de, possibilitar a construção dos modelos; como também, consiste na disponibilização das métricas de avaliação do modelo para avaliação da qualidade. Portanto, o procedimento por meio da utilização deste programa seguiu o mesmo padrão da modelagem por homologia, com as seguintes etapas: 1) seleção do modelo de MrACBP utilizando a sequência FASTA da proteína de ligação à acil-CoA (ACBP) que foi obtida pelo banco de dados do NCBI (https://www.ncbi.nlm.nih.gov/protein/554912798) do fungo *Moniliophthora roreri* MCA 2997 (ESK94037.1) (MEINHARDT et al., 2014); 2) alinhamento do modelo alvo MrACBP com o provável molde MpACBP proposta por Monzani et al. (2010), considerando o alinhamento de boa qualidade cobertura de e identidade acima de 30%; 3) sendo satisfatória para construção do modelo tridimensional (3D) da ACBP de *M.*

roreri; e 4) na etapa de avaliação; após a construção do modelo, foi necessário identificarmos possíveis erros relacionados a escolha da estrutura de referência, e ao alinhamento entre a sequência-alvo de MrACBP e o provável molde MpACBP). A literatura considera o modelo de boa qualidade quando apresentados os valores de GMQE (*Global Model Quality Estimation*) que se refere a estimativa de qualidade do modelo, expressa a pontuação de estimativa global da estrutura, avaliada entre 0 para modelos não válidos e 1 para modelos excelentes e o valor apresentado do QMEAN que fornece estimativas de qualidade absoluta globais (para toda estrutura), e locais (por resíduos), sendo considerado válido acima de -4.0. Os quatro termos individuais das pontuações de qualidade globais da QMEAN (valores numéricos próximos de zero) indicam que a propriedade é semelhante ao que se esperaria de estruturas experimentais de tamanho semelhante, portanto, valores positivos indicam que os *scores* do modelo são superiores às estruturas experimentais em média (SCHWEDE et al., 2003; MANDHAVI SASTRY et al., 2013; WATERHOUSE et al., 2018). De modo que estas etapas possam ser repetidas até que a estrutura do modelo da proteína final fosse satisfatória (WATERHOUSE et al., 2018; SCHWEDE et al., 2003; RANGEL, 2019).

4.3 Minimização de energia e validação da ACBP de *Moniliophthora roreri*

Em seguida, a estrutura proteica da ACBP foi refinada e realizado a minimização de energia através do pacote AMBER 20 (CASE et al. 2020) e validada a qualidade através da variação no Desvio da Raiz Média Quadrática com a sigla em inglês RMSD (Root Mean Square) (MADHAVI SASTRY et al., 2013) de aproximadamente 1Å, quando se compararam as sequências da estrutura experimental que serviu de molde e o modelo obtido, usando o programa PyMOL 2.5.2 (SCHRÖDINGER, 2021) para o alinhamento das sequências de aminoácidos e comparação das estruturas 3D da proteína ACBP de *M. roreri* e a proteína homóloga ACBP de *Moniliophthora perniciosa* (MONZANI et al., 2010).

O PROCHECK que valida a qualidade do modelo construído e avalia diversos parâmetros estereoquímicos de estruturas de proteínas, de importância fundamental como os comprimentos de ligação, os ângulos planos, a planaridade dos anéis de cadeias laterais, os ângulos torcionais da cadeia principal, por isso foi utilizado gerando o gráfico de Ramachandran, para definirmos os resíduos que se encontravam nas regiões conformacionalmente mais favoráveis e desfavoráveis, além de fornecera avaliação da qualidade do modelo-alvo (LASKOWSI, R.A.*et al.*, 1993).

4.4 Triagem virtual baseada no receptor (RBVS)

A partir da busca em revisão bibliográfica, foram encontrados 28 compostos (FRITZLER & ZHU, 2011) e outros 5 compostos semelhantes e moléculas ativas depositadas em bancos de dados de compostos Drug Bank (https://go.drugBank.com); a partir destes foram considerados as estruturas de 24 candidatos à inibidores ou ligantes conhecidos da ACBP em outros organismos, onde posteriormente foram catalogados os respectivos nomes, código de identificação (ID), fórmula molecular e descritor simplificado SMILES canônico, de acordo com o registro da estrutura química no banco de dados de acesso público PubChem (https://pubchem.ncbi.nlm.nih.gov). Em seguida o programa MarvinSketch 19.25 (https://chemaxon.com/products/marvin) foi utilizado para desenhar as estruturas 3D dos compostos selecionados e salvando nos formatos mol2 e pdb.

4.5 Preparação do receptor e ligante para o acoplamento molecular (*docking*)

Para realizarmos o *docking* molecular, as moléculas dos ligantes e receptor foram preparadas utilizando o software MarvinSketch 19.25 (https://chemaxon.com/products/marvin) para conversão dos ligantes de SMILES canônico das moléculas para estrutura 3D e correção dos possíveis erros de torção e salvamos em formato mol2 e pdb. Utilizamos separadamente o arquivo mol2 do receptor e ligantes, que serviu como input para ser convertido em formato pdbqt no AutoDock Tools versão 1.5.6 (MORRIS et al., 2009) e selecionamos os parâmetros do grid box (figura 22), de modo a envolver toda a macromolécula, de acordo o posicionamento do seu sítio ativo, que foram desenhadas de acordo com o sítio da ACBP de *M. perniciosa* homóloga resolvida por cristalografia descrito por Monzani et al. (2010); além de, adicionarmos átomos de hidrogênio ao receptor para computar o seu estado de protonação.

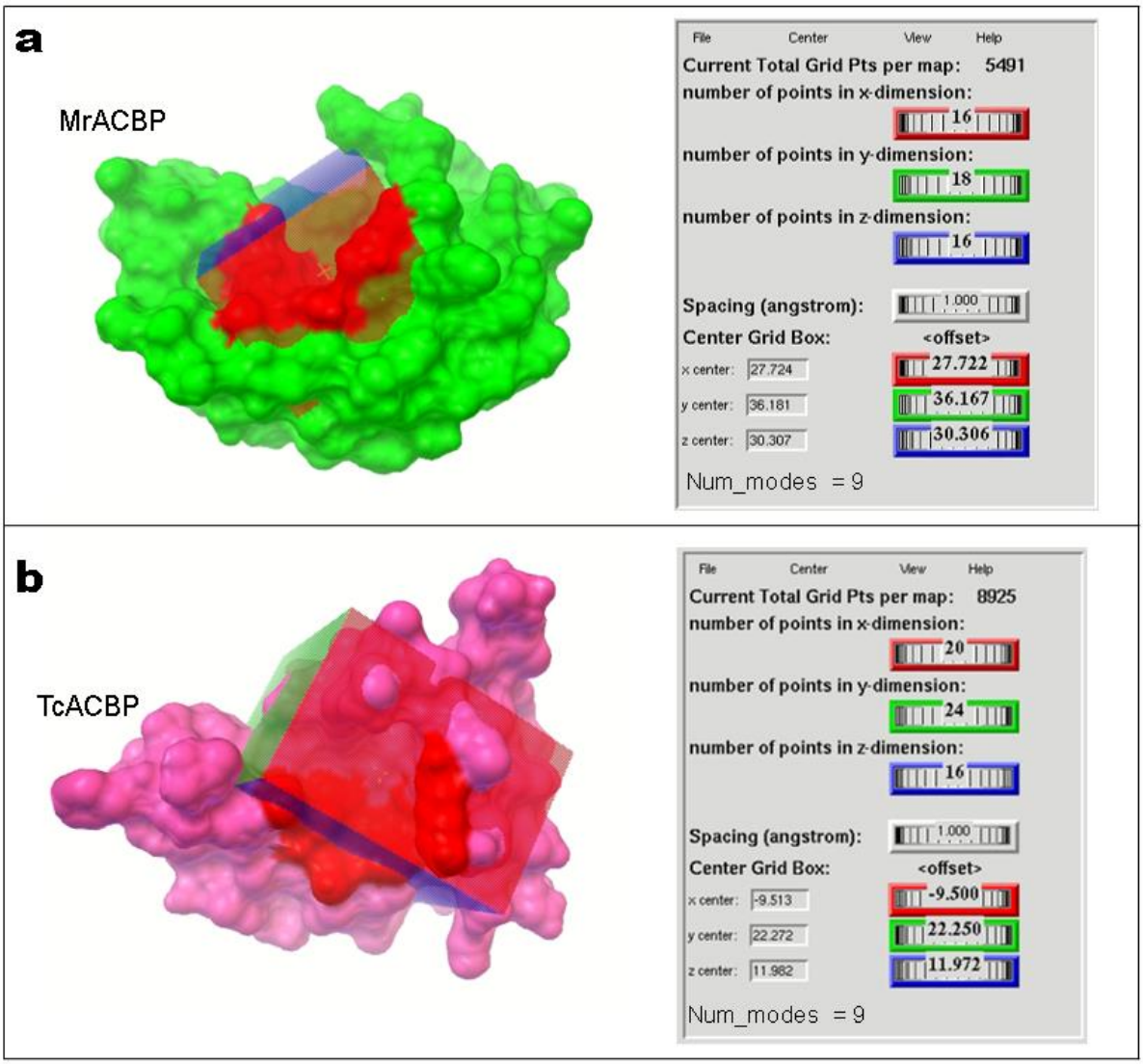

Figura 22: Representação da construção das coordenadas de x, y e z do sítio ativo do receptor (a) MrACBP e (b)TcACBP para realizar o acoplamento molecular (*docking*). Na região central da proteína mostra a caixa (*grid box*) no meio do bolso de ligação. Ao lado direito da imagem as configurações das coordenadas de x, y e z para ser utilizado no AutoDock Vina, como local de interação com ligantes. Num_modes: se refere a quantidade de graus de liberdade, neste caso 9 *poses*, para as ligações intermoleculares. **Fonte:** Imagem gerada no AutoDock Tools v.1.5.6; MORRIS et al., (2009).

4.6 Processo de Acoplamento molecular *(Docking)*

Para o cálculo de *docking* molecular usamos o AutoDock Vina (TROTT; OLSON, 2010), onde foram geradas 9 posições (*poses*) de *docking* para cada ligante com o sítio ativo (MONZANI et al., 2010) da ACBP de *M. roreri* e também; realizamos o *docking* com a ACBP de *T. cacao*; para avaliarmos as interações destes ligantes e seus valores de energia de afinidade. As posições do *docking* foram avaliadas usando o PyMOL 2.5.2 (SCHRÖDINGER, 2021) para selecionarmos a melhor posição (*pose*) de cada ligante com o sítio da ACBP de *M. roreri*, e com o sítio da ACBP de *T. cacao*. Para isso foi considerada a energia de afinidade entre o

ligante e o receptor, bem como a interação dos ligantes com os aminoácidos do sítio ativo e o desvio médio quadrado (RMSD) (MADHAVI SASTRY et al., 2013). Por fim, analisamos as interações entre receptor-ligante dos melhores ligantes na construção do mapas de interação 3D através da ferramenta PLIP Web Tool (*Protein-Ligand Interaction Profiler*) (https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index) (ADASME et al., 2021).

4.7 Modelagem por homologia da ACBP de *Theobroma cacao*

A estrutura tridimensional (3D) da ACBP de *Theobroma cacao* foi construída por meio da modelagem por homologia, a partir da busca por sequências de aminoácidos depositada no NCBI (https://www.ncbi.nlm.nih.gov/protein), apresentando 13 sequências disponíveis de TcACBP. Posteriormente, utilizando o BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi), restrito à sequências com estruturas depositadas no *Protein Data Bank* (PDB) (https://www.rcsb.org/), buscamos por sequências de proteínas homólogas com estrutura tridimensional resolvida por cristalografia ou ressonância magnética nuclear considerando o maior percentual de cobertura e identidade com a sequência em estudo, portanto, optamos pela sequência que apresentou mais relevante para construção do modelo de TcACBP a sequência com os resultados do BLASTP um melhor alinhamento, indicando-a como provável molde, com cobertura (valor do *score* positivamente correlacionado com a qualidade do alinhamento, ou seja, quanto maior este valor, melhor o alinhamento) e identidade (acima de 30%) entre os resíduos de aminoácidos. Esses *scores* foram analisados utilizando o programa de alinhamentos múltiplos de sequências de proteínas, o T-COFFEE (http://tcoffee.crg.cat/) (DI TOMMASO et al., 2011).

Em seguida, a proteína ACBP de *Oryza sativa* de acesso PDB ID 5H3G (https://www.rcsb.org/structure/5H3G) foi utilizada como molde para a construção do modelo 3D da TcACBP. Na construção do modelo utilizamos o servidor SWISS-MODEL Workspace (http://swissmodel.expasy.org) (WATERHOUSE et al., 2018; SCHWEDE et al., 2003). Finalmente, a estrutura proteica da TcACBP foi minimizada através do servidor online Galaxy WEB (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) (KO, J. et al., 2012) e validada a qualidade através da variação no Desvio da Raiz Média Quadrática (RMSD) (MADHAVI SASTRY et al., 2013) usando o programa PyMOL 2.5.2 (SCHRÖDINGER, 2021).

4.8 Simulações de Dinâmica Molecular

Realizamos as simulações de Dinâmica Molecular de energia livre dos melhores complexos selecionados na etapa de acoplamento (*docking*), utilizando o software de código aberto GROMACS 2020 (*Protein-Ligand Complex*) (http://www.mdtutorials.com/gmx/complex/index.html) (ABRAHAM, et al., 2015; LEMKUL, 2018). A previsão e refinamento prévio das proteínas MrACBP e TcACBP foram realizados através do servidor GalaxyWeb (https://galaxy.seoklab.org/) (KO et al., 2012). Os processos de minimização de energia, parametrização da MrACBP e TcACBP, como também a dinâmica molecular de estabilidade entre MrACBP-ligantes, foram realizados utilizando o campo de força CHARMM (BROOKS et al., 1983) com a finalidade de estabelecer as interações e distâncias dos átomos ligados e não ligados; que foi construído a partir de um pacote de programa especializado em simulação e otimização de biomoléculas e sistemas orgânicos, o CGenFF 2018 (https://cgenff.umaryland.edu/) (MADHAVI SASTRY et al., 2013).

Utilizando o programa "CGenFF", geramos os arquivos de topologia com informações de tipos de átomos, cargas, conectividade vinculada, coordenadas, solvatação da proteína estudada. Realizamos adição de íons na parametrização (Na^+ ou Cl^- de acordo com a carga da proteína) afim de neutralizar o sistema. Logo, utilizamos o programa AVOGADRO 1.1.1 (https://sourceforge.net/projects/avogadro/) (HANWELL *et al.*, 2012) que possibilitou adicionarmos os átomos de hidrogênio, uma vez que, estruturas de cristal normalmente não atribuem coordenada de hidrogênio, então elas devem ser incorporadas. No servidor Web CHARMM-GUI (https://www.charmm-gui.org/) (JO et al., 2008; JO et al., 2016) realizamos a parametrização dos complexos proteína MrACBP e TcACBP com os respectivos ligantes

No processo seguinte, seguindo o tutorial do GROMACS 2020 (*Protein-Ligand Complex*) (http://www.mdtutorials.com/gmx/complex/index.html) (ABRAHAM et al., 2015; LEMKUL, 2018), calculamos as forças exercidas sobre cada partícula devido as interações interatômicas utilizamos o campo de força CHARMM 2020 (HUANG et al., 2016); em seguida realizamos a otimização da estrutura, e as simulações da dinâmica da estrutura através da integração da equação de Newton por métodos numéricos e por fim; foram analisados os resultados através das propriedades de equilíbrio; quem em tese nos permitirá compreender a estrutura energética e a dinâmica molecular do sistema de interação da proteína com os prováveis ligantes (LEMKUL, 2018).

5. RESULTADOS E DISCUSSÃO

5.1 Análise da sequência da ACBP de *Moniliophthora roreri* (MrACBP)

A sequência FASTA da proteína de ligação à acil-CoA (ACBP) foi obtida pelo banco de dados do NCBI (https://www.ncbi.nlm.nih.gov/protein/554912798) do fungo *Moniliophthora roreri* MCA 2997 (ESK94037.1) que teve seu genoma sequenciado e identificado no trabalho de Meinhardt et al. (2014). A análise desta sequência mostrou que ACBP codifica uma sequência de 104 aminoácidos com ponto isoelétrico e massa molecular de 5.52 e 11854.48 kDa, respectivamente, utilizando o Compute Pi/Mw Expasy (https://web.expasy.org/compute_pi/) (GASTEIGER et al., 2005). A análise da sequência no servidor SignalP-5.0 (http://www.cbs.dtu.dk/services/SignalP/) (PETERSEN et al., 2011) apresentou a estimativa da probabilidade de ocorrência do peptídeo sinal foi de 0,001; portanto, não indicou a presença do peptídeo sinal, como também não mostrou a localização do sítio de clivagem entre os aminoácidos, e nem de regiões desordenadas na sequência de aminoácidos da proteína de interesse (figura 23).

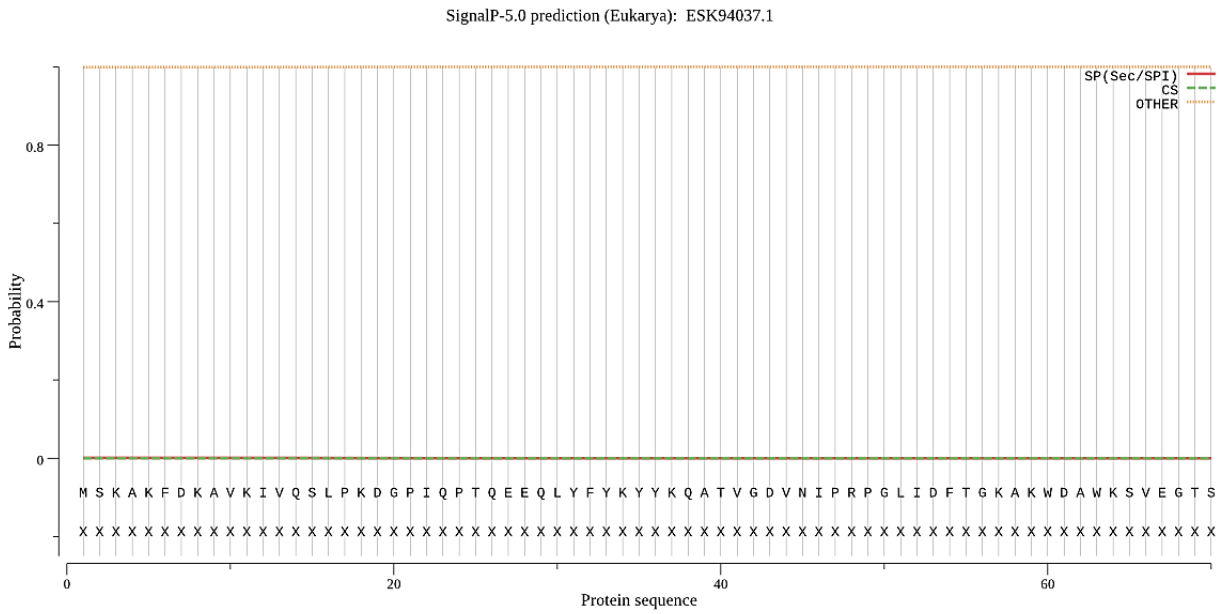

Figura 23: Resultado da predição do peptídeo sinal da proteína MrACBP. No eixo horizontal está representado a sequência de aminoácidos que constituem a proteína e no eixo vertical a estimativa da probabilidade de ocorrência do peptídeo sinal, que foi de 0,001 ou 0,1%. SP: peptídeo sinal (Sec/SPI) corresponde ao traço vermelho. **Fonte:** Predição SignalP -5.0.

A predição realizada no banco de dados de domínios conservados (CDD) (MARCHLER-BAUER et al., 2014) no NCBI na busca por domínios conservados na sequência desta proteína (figura 24) resultou no domínio conversado ACBP (número de acesso: pfam00887) entre os aminoácidos 3 até o 87, com Bit Score de 109,19 (se refere a pontuação de qualidade do alinhamento entre as sequências, quanto maior esse valor mais significativo é a correspondência do alinhamento), e *E-value* de 7,63e-33 (quanto menor o valor de *E-value*,

menor a probabilidade da sequência de correspondência seja encontrada ao acaso em banco de dados). As unidades são conservadas evolutivamente sendo pertencente à super família ACBP, classificando como proteína de ligação a acil-CoA que se liga a ésteres de acil-CoA e também pode atuar desempenhando um papel na manutenção e/ou tráfego. No InterPro também identificou o domínio conservado ACBP, e através deste servidor mostrou que no banco de dados PROFILE identificou o domínio ACB, em ambos iniciando no aminoácido 2 até o 91, também são membros da super família ACBP.

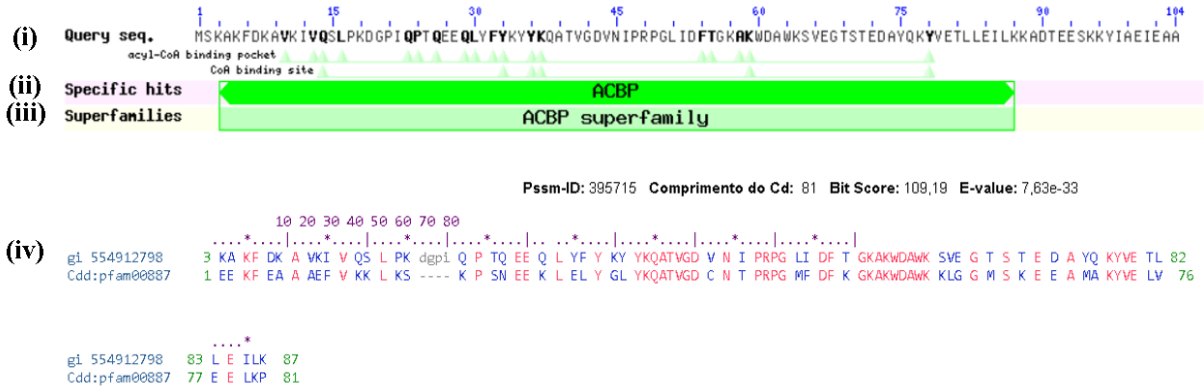

Figura 24: Domínio e superfamília da proteína ACBP de *M. roreri* (NCBI ID - ESK94037.1). (i) Sequência de aminoácidos da MrACBP, em "negrito" os resíduos funcionais, e os triâncgulos apontam os resíduos correspondentes a região de ligação e o sítio ativo. (ii) Domínio conservado ACBP dos resídios 2 até o 81, com "hits" de maior pontuação desta sequência. (iii) Super família a qual pertence o domínio ACBP. (iv) Alinhamento entre a sequência MrACBP (primeira linha) e domínio conservado ACBP - Cdd: pfam00887(segunda linha), em destaque de cor vermelha resíduos de aminoácidos idênticos. Pssm-ID:395715 (identificador do domínio na posição específica na matriz de pontuação). Comprimento do Cd (domínio conservado): 81resíduos de aminoácidos, Bit Score: 109,19 (qualidade do alinhamento), E-value: 7,63e-33 (probabilidade da sequência de correspondência seja encontrada ao acaso) **Fonte:** Predição CDD, MARCHLER-BAUER et al. (2014)

5.2 Modelagem da estrutura 3D de MrACBP

Em princípio, foi realizada a busca no banco de dados do NCBI por sequências de proteínas de ligação acil-CoA (ACBP) do fungo basidiomiceto *Moniliophthora roreri* para ser utilizada como proteína alvo. Nesta etapa, foram encontradas 6 sequências (tabela 2), das quais foram excluídas as anotações que não possuíssem um tamanho maior do que 104 resíduos de aminoácidos, sobrando duas anotações, das quais excluímos a proteína hipotética devido a falta de evidência experimental de que é expressa *in vivo*, e optamos pela sequência ESK94037.1.

Tabela 2 - Resultados da busca por ACBP de *Moniliophthora roreri* no banco de dados de proteínas (NCBI) e pesquisa de alinhamento local por estruturas homólogas na ferramenta do BLASTP. Em destaque "cinza" sequência ESK94037.1 referente a MrACBP de entrada com melhor alinhamento local com o molde MpACBP.

ACBP de *Moniliophthora roreri* no NCBI	Número de aminoácidos na sequência	Estrutura proteica de melhor alinhamento no BLASTP*
proteína hipotética WG66_10451 [*Moniliophthora roreri*] (KTB36995.1)	455 aa	PDB ID: 3FP5 (ACBP de *M. perniciosa*) Identidade[a] 45.54% Cobertura[b] 21%
proteína hipotética WG66_12087 [*Moniliophthora roreri*] (KTB35489.1)	1423 aa	PDB ID: 6GAQ (Estrutura do complexo de ligase succinil-CoA formador de ADP humano SUCLG1-SUCLA2) Identidade[a] 64.26% Cobertura[b] 21%
proteína hipotética WG66_19578 [*Moniliophthora roreri*] (KTB27791.1)	104 aa	PDB ID: 3FP5 (ACBP de *M. perniciosa*) Identidade[a] 91,35% Cobertura[b] 100%
proteína de ligação de acil [*Moniliophthora roreri* MCA 2997] (ESK97986.1)	230	PDB ID: 4GMR (Estrutura de cristal da proteína projetada) Identidade[a] 51,19% Cobertura[b] 36%
Proteína de ligação a acila [*Moniliophthora roreri* MCA 2997] (ESK94037.1)	104 aa	PDB ID: 3FP5 (ACBP de *M. perniciosa*) Identidade[a] 91,35% Cobertura 100%
transportador de ácidos graxos de cadeia longa [*Moniliophthora roreri* MCA 2997] (ESK93709.1)	457 aa	PDB ID: 3FP5 (ACBP de *M. perniciosa*) Identidade[a] 46,53% Cobertura[b] 22%

Fonte: Adaptado de Predição BLASTP (NCBI) disponível em https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins . Desenvolvido por ALTSCHUL et al. (1990); ALTSCHUL et al. (1997)
*Estruturas que apresentaram os melhores valores de similaridades entre as sequências de acordo com a ACBP de *M. roreri* da coluna 1, alinhamento realizado na ferramenta BLASTP.
[a] Refere-se a fração de aminoácidos idênticos entre pares de sequências após um alinhamento dessas sequências
[b] Refere-se a cobertura da região das sequências de aminoácidos das estruturas que foram alinhadas

Resultados do BLASTP (tabela 2) indicaram um melhor alinhamento com a sequência da ACBP do fungo *Moniliophthora perniciosa* proposta por Monzani et al. (2010), indicando-a como provável molde, onde o alinhamento com a sequência ESK94037.1 apresentou cobertura de 100% e identidade de 91,35%, sendo satisfatória para construção do modelo tridimensional (3D) para a construção da ACBP de *M. roreri*. A proteína alvo foi modelada a

partir do alinhamento da sequência alvo com a estrutura tridimensional 3FP5 da ACBP de *M. perniciosa* (MpACBP) (Figura 25), com resolução cristalográfica de 1.61 Å e *R-value* de 0.186.

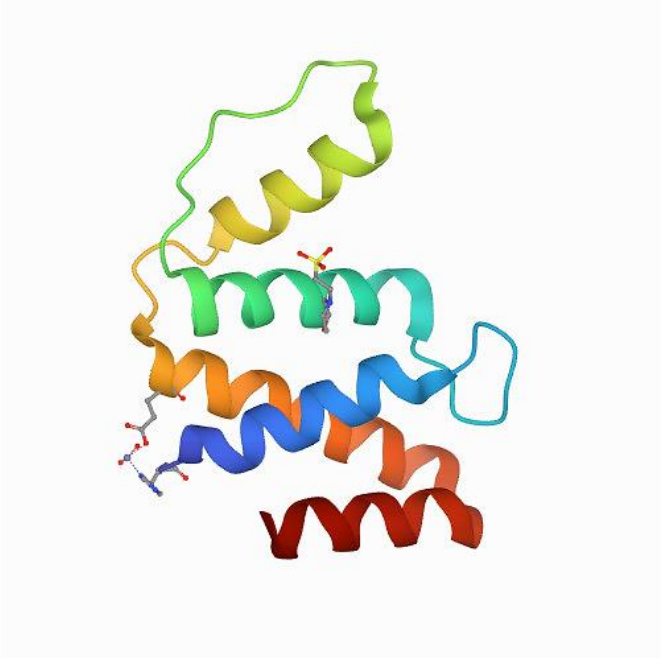

Figura 25: Estrutura tridimensional da ACBP de *Moniliophthora perniciosa*. Fonte: RCSB PDB 3FP5.

Antecipando o processo de construção do modelo, utilizamos o servidor SWISS-MODEL Workspace, para também procurarmos modelos homólogos (moldes similares) a partir da sequência alvo (sequência FASTA determinada anteriormente para *M. roreri*). Na figura 26, (i) mostra o resultado gerado no servidor que apontou o mesmo molde encontrado no PDB, com 91,26% de identidade entre os resíduos, sendo que para modelos de boa qualidade consideram-se valores acima de 30% de identidade entre a proteína-molde e a proteína-alvo; 58% grau de similaridades entre as sequências de aminoácidos e cobertura de 99%. Antes da localização da estrutura primária, o programa automaticamente realizou o alinhamento sequencial entre o alvo e o molde a fim de encontrar a porcentagem de similaridade residual. A figura 26, tirada da página do SWISS-MODEL Workspace apresenta o (ii) alinhamento de aminoácidos (WATERHOUSE et al., 2018; SCHWEDE et al., 2003)

i

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Range	Coverage	Description
3fp5.1.A	91.26	monomer	0.00	HHblits	X-ray	1.61Å	0.58	2 - 104	0.99	Acyl-CoA Binding Protein

ii

```
Target    MSKAKFDKAVKIVQSLPKDGPIQPTQEEQLYFYKYYKQATVGDVNIPRPGLIDFTGKAKWDAWKSVEGTSTEDAYQKYVE
3fp5.1.A  -SKAKFDKAVEIVQSLPKDGPIKPTQDEQLYFYKYFKQATVGDVNISRPGLMDFTGKAKWDAWKSVEGTSKEVAYQKYVE
```

iii

Ligand Name.Number	Reason for Exclusion	Description
MES.2	Binding site not conserved.	2-(N-MORPHOLINO)-ETHANESULFONIC ACID
ZN.1	Binding site not conserved.	ZINC ION

Figura 26: Resultado da busca de modelos homólogos no SWISS MODEL. (i) Modelo 3fp5.1.A com 91,26% de identidade, 0.58 similaridades entre as sequências de aminoácidos e cobertura de 0.99; (ii) alinhamento das sequências da proteína-alvo com o molde homólogo; (iii) presença da molécula MES e ZN no sítio de ligação do molde homólogo. **Fonte:** Predição SWISS-MODEL Workspace.

Na construção do modelo, observamos que a estrutura tridimensional (3D) da ACBP de *M. roreri* é formada por cinco alfa-hélices com quatro regiões de alças (loops) ilustrado na figura 27; diferente da topologia típica das ACBPs que apresentam um arranjo único entre quatro dobras em hélice (H1-H4) e em semelhança estrutural com a ACBP de *M. perniciosa* (MONZANI et al., 2010) que constitui também de uma quinta hélice em sua topologia.

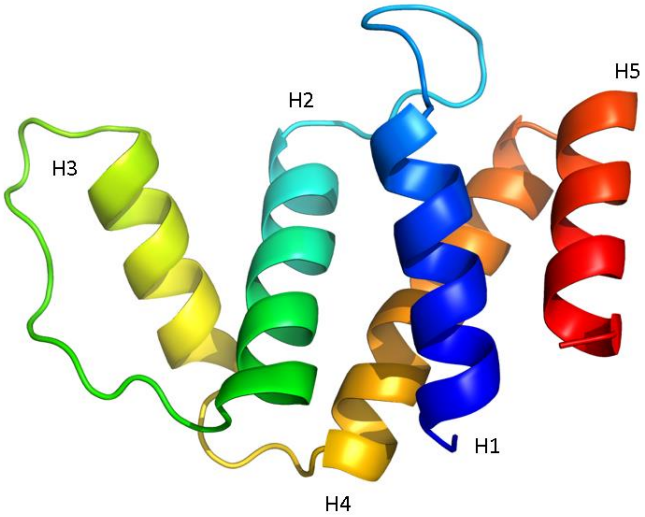

Figura 27: Estrutura tridimensional da ACBP de *Moniliophthora roreri* (modelo alvo). Na representação de fita colorida, apresentando cinco alfa-hélices (H1-H5) rotuladas, construída a partir do molde MpACBP (*M. perniciosa*). **Fonte:** Imagem adaptada gerada no PyMol; SCHRÖDINGER (2021)

No alinhamento estrutural da MrACBP com energia minimizada (CASE et al. 2020) e MpACBP, em que utilizamos o programa PyMOL 2.5.2 (SCHRÖDINGER, 2021), o RMSD (derivação da raiz quadrada média) foi de 0.469 Å , valor que representa a existência de

mínimas diferenças entre o molde-modelo, ilustrado na figura 28 (MADHAVI SASTRY et al., 2013).

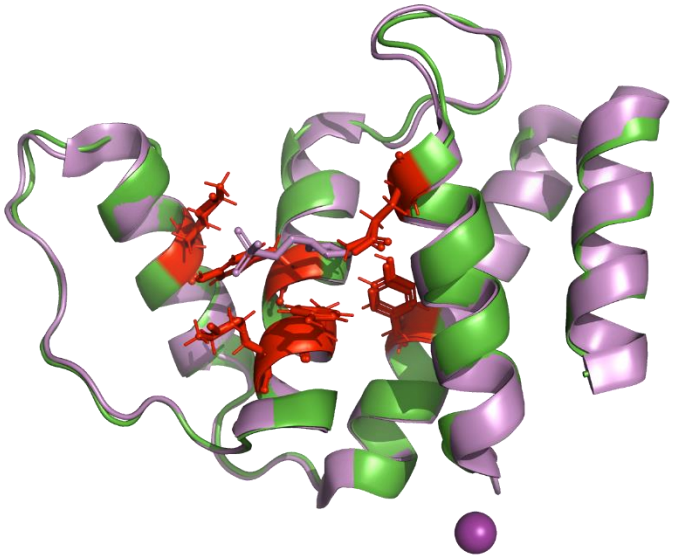

Figura 28: Alinhamento estrutural entre MrACBP com energia minimizada – modelo construído (verde) e MpACBP (PDB 3FP5) - molde (lilás). Incluindo a região do sítio em destaque (vermelho) mostrando mínimas divergências entre o molde-modelo. A esfera na cor roxa mostra íon zinco (ZN^{2+}). **Fonte:** Imagem gerada pelo programa PyMOL; SCHRÖDINGER (2021).

Na figura 29, representa o alinhamento realizado no T-COFFEE demonstrou que entre a ACBP de *M. roreri* e seu molde cristalográfico de código PDB 3FP5 o SCORE (qualidade do alinhamento estre as sequências) de 100% e 99% de cobertura, mostrando os resíduos conservados entre o modelo e o seu molde destacado (rosa) (DI TOMMASO et al., 2011).

```
        T-COFFEE, Version_11.00 (Version_11.00)
        Cedric Notredame
        CPU TIME:0 sec.
        SCORE=100
        *
         BAD AVG GOOD
        *
        ESK94037.1     : 100
        pdb|3FP5|A     : 100
        cons           :  10

MrACBP  ESK94037.1     --MSKAKFDKAVKIVQSLPKDGPIQPTQEEQLYFYKYYKQATVGDVNIPRPGLIDFTGKAKWDAWKSVE
MpACBP  pdb|3FP5|A     SHMSKAKFDKAVEIVQSLPKDGPIKPTQDEQLYFYKYFKQATVGDVNISRPGLMDFTGKAKWDAWKSVE

        cons             **********:***********:***:********:**********.****:***************


MrACBP  ESK94037.1     GTSTEDAYQKYVETLLEILKKADTEESKKYIAEIEAA
MpACBP  pdb|3FP5|A     GTSKEVAYQKYVEKLLEILKKADTEESKKYIAEIEAA

        cons           ***.* *******.************************
```

Figura 29: Alinhamento entre as sequências de aminoácidos da MrACBP e MpACBP (PDB 3FP5). Fonte: Predição gerada no T-COFFEE, adaptado.

Após a construção do modelo, foi necessário identificarmos possíveis erros relacionados a escolha da estrutura de referência, e ao alinhamento entre a sequência-alvo e molde. Na figura 30, o GMQE (*Global Model Quality Estimation*) (1) se refere a estimativa de qualidade do modelo, expressa a pontuação de estimativa global da estrutura, avaliada entre 0 para modelos não válidos e 1 para modelos excelentes, o valor obtido no servidor foi de 0.93 (WATERHOUSE et al., 2018; SCHWEDE et al., 2003). A literatura considera o modelo de boa qualidade, uma vez que, o valor apresentado do QMEAN que fornece estimativas de qualidade absoluta globais (para toda estrutura), e locais (por resíduos) apresentado foi de 1.50, sendo considerado válido acima de -4.0. Os quatro termos individuais das pontuações de qualidade globais da QMEAN também são listados em (3), a área branca nos gráficos de barras (valores numéricos próximos de zero) indica que a propriedade é semelhante ao que se esperaria de estruturas experimentais de tamanho semelhante, portanto, os valores positivos apresentados indicam que os *scores* do modelo são superiores às estruturas experimentais em média (SCHWEDE et al., 2003; MANDHAVI SASTRY et al., 2013; WATERHOUSE et al., 2018).

Na figura 30, o gráfico "Qualidade Local" (4) mostra, para cada resíduo do modelo (relatado no eixo x), a semelhança esperada com a estrutura nativa (eixo y), neste modelo os resíduos se concentram acima da pontuação de 0.6, indicando boa qualidade. No gráfico "Comparação" (5), as pontuações de qualidade do modelo construído são comparadas com pontuações obtidas por estruturas experimentais de tamanho similar, o eixo x mostra o comprimento da proteína (número de resíduos), e o eixo y é a pontuação QMEAN normalizada, indicando que o modelo atual é representado com uma estrela vermelha, os pontos pretos são estruturas experimentais com um score QMEAN normalizado dentro de 1 desvio padrão da média (| Z-score | entre 0 e 1), estruturas experimentais com um | Z-score | entre 1 e 2 são cinza claro. A média e o desvio padrão das estruturas experimentais em torno da localização x da estrela são a base para calcular a pontuação Z QMEAN do modelo (2), ou seja, quantos desvios padrão das pontuações médias modelo (SCHWEDE et al., 2003; MANDHAVI SASTRY et al., 2013; WATERHOUSE et al., 2018).

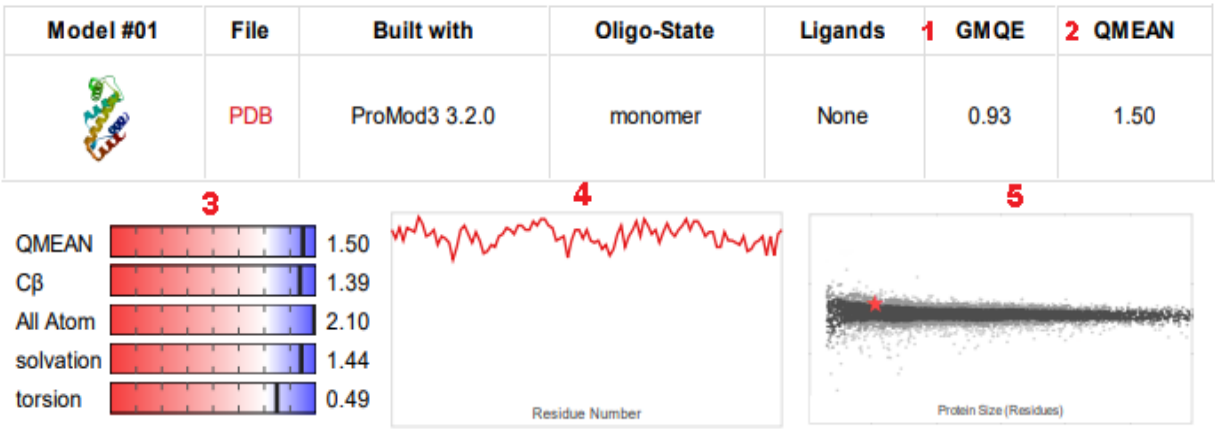

Figura 30: Esquema de avaliação e validação do modelo homólogo PDB 3FP5. Em (1) GMQE (estimativa de qualidade do modelo) de 0.93; (2) QMEAN (estimativa de qualidade absoluta globais e locais) de 1.50; (3) quatros termos individuais das pontuações de qualidade globais da QMEAN com valores positivos; (4) gráfico "Qualidade Local"; (5) gráfico "comparação". **Fonte:** Predição SWISS-MODEL Workspace, adaptado.

O servidor PROCHECK gerou os gráficos de Ramachandran (figura 31), indicando que o modelo construído se encontra com 94, 6% dos resíduos da estrutura alvo localizados em zonas energeticamente mais favoráveis (áreas em vermelho), e 5,4% em regiões favoráveis (áreas amarelas), considerando-se a análise de 118 estruturas com uma resolução de pelo menos 2.0 Å e R-factor não maior que 20%, sendo, portanto, o modelo de boa qualidade (acima de 90%) (LASKOWSI et al., 1993).

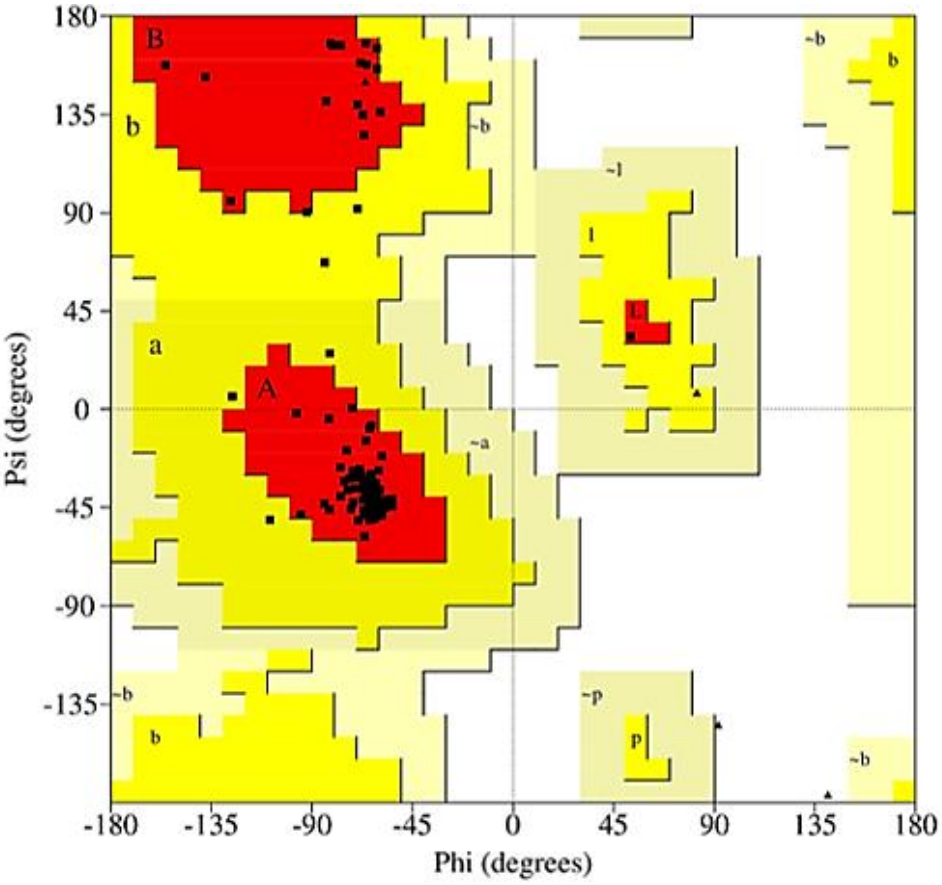

Figura 31: Gráfico de Ramachandran, indicando boa qualidade do modelo construído (acima de 90%). As regiões apresentadas em vermelho, são as mais favoráveis, as representadas em amarelo, são as favoráveis, as representadas em bege, são as menos favoráveis e as regiões brancas, são as desfavoráveis. **Fonte:** Predição PROCHECK.

A região do sítio ativo da ACBP foi descrita por Monzani et al. (2010) para *M. perniciosa* (GLN 14, TYR 33, PHE 36, LYS 37, LYS 59, TYR 78), porém, realizamos o alinhamento estrutural no PyMOL 2.5.2 (SCHRÖDINGER, 2021) com os resíduos de aminoácidos da ACBP de *M. roreri*, e o sítio ativo ficou definido para MrACBP composto pelos resíduos de aminoácidos: GLN 14, TYR 33, TYR 36, LYS 37, LYS 59, TYR 78 (figura 32 a). O sítio da MrACBP mostrou-se com o aminoácido TYR 36 em vez de PHE 36, os demais estão de acordo com o sítio de MpACBP. Além disso, o sítio de MrACBP apresentou a maior concentração de aminoácidos polares neutros sendo estes os aminoácidos: GLN 14, TYR 33, TYR 36, TYR 78, este grupo de aminoácidos possuem cadeias laterais polares eletricamente neutras em pH neutro, e o grupo hidroxila na tirosina é ligado a um grupo hidrocarboneto aromático, o qual perde um próton em meio alcalino (pH > 10) e forma a base conjugada. Os demais que compõem o sítio ativo incluem os aminoácidos polares básicos: LYS 37, LYS 59 que possuem cadeias laterais básicas e carregada positivamente em pH neutro. A partir da

análise da superfície molecular da MrACBP podemos visualizar uma extensa cavidade (hidrofóbica) localizada na região de interação com ligantes, do sítio ativo (figura 32 b).

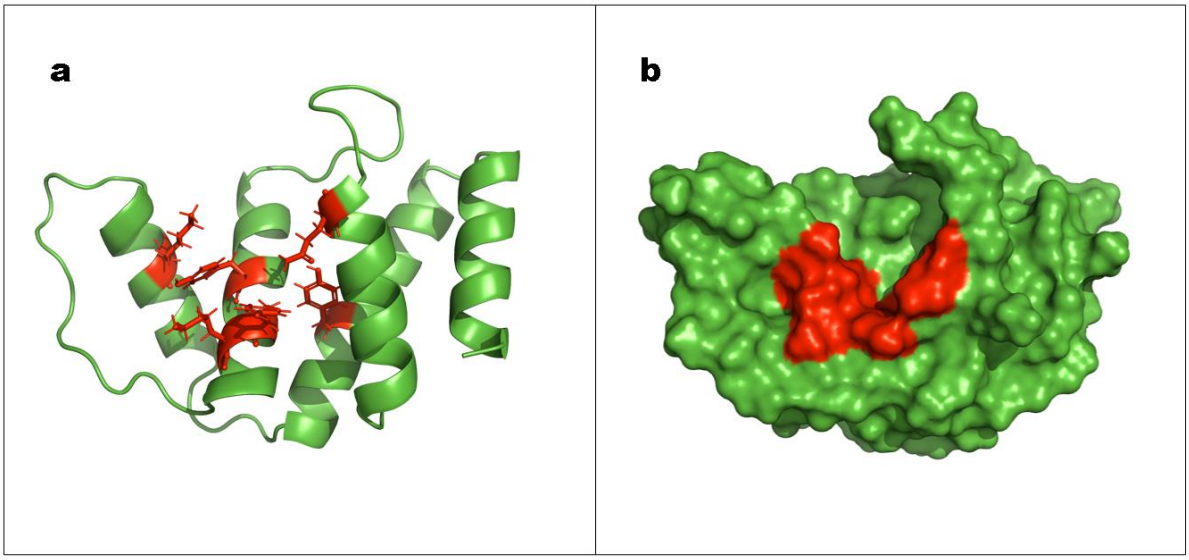

Figura 32: a e b. **Representação estrutural (3D) da MrACBP (verde) com região do sítio ativo destacado (vermelho) composto pelos resíduos de aminoácidos: GLN 14, TYR 33, TYR 36, LYS 37, LYS 59, TYR 78.** Em b detalhando sua cavidade hidrofóbica na região do sítio ativo. **Fonte:** Imagem gerada pelo PyMOL; SCHRÖDINGER (2021).

5.3 Modelagem da estrutura 3D da ACBP de *Theobroma cacao* (TcACBP)

Inicialmente, buscamos no banco de dados do NCBI por sequências de proteínas de ligação acil-CoA (ACBP) de *Theobroma cacao*, das 13 sequências disponíveis, portanto, optamos pela sequência EOY25439.1 que se refere a proteína 6 de ligação a acil-CoA isoforma 2 de *T. cacao* que apresentou nos resultados do BLASTP um melhor alinhamento, com a sequência da proteína 2 de ligação a acil-CoA de *Oryza sativa* (OsACBP) (GUO et al., 2017)., indicando-a como provável molde, com cobertura de 75% e identidade de 83,95% entre os resíduos de aminoácidos. Realizamos a construção do modelo tridimensional (3D) da proteína-alvo TcACBP, utilizando o servidor do SWISS-MODEL Workspace (http://swissmodel.expasy.org) (WATERHOUSE et al., 2018; SCHWEDE et al., 2003), que foi modelada a partir do alinhamento da sequência alvo com a estrutura cristalográfica de código PDB 5H3I da OsACBP, a estrutura da TcACBP, ilustrado na figura 33, é formada por quatro alfa-hélices com regiões de três alças (loops) conectando-as e; com cobertura de 84% e

identidade 81,11% entre os resíduos de aminoácidos; além de, resolução cristalográfica de 2.30 Å e *R-value* de 0.223.

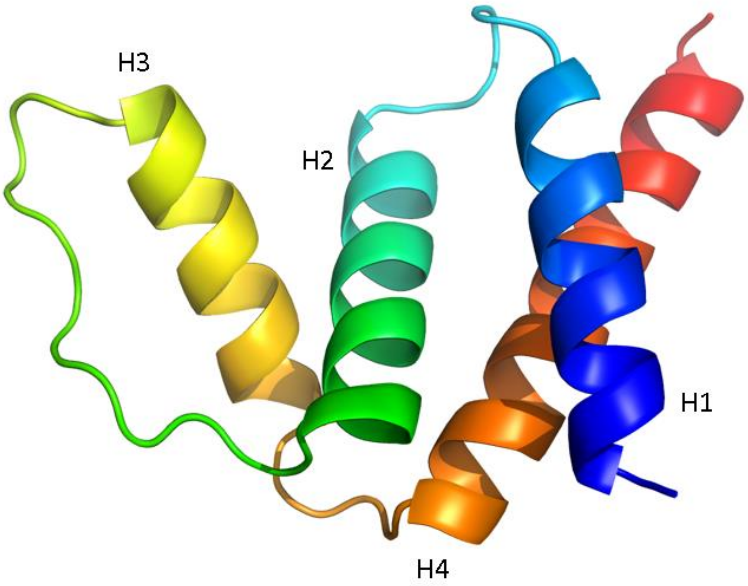

Figura 33: Estrutura tridimensional de TcACBP (*Theobroma cacao*). Na representação de fita colorida, apresentando as quatro alfa-hélices (H1-H4) rotuladas; construída a partir do molde OsACBP (*Oryza sativa*). **Fonte:** Imagem gerada no PyMol; SCHRÖDINGER (2021).

No alinhamento feito no T-COFFEE (DI TOMMASO et al., 2011) (figura 34) mostrou SCORE (qualidade do alinhamento estre as sequências) de 100% entre a ACBP de *Theobroma cacao* (TcACBP) e seu molde cristalográfico ACBP de *Oryza sativa* (OsACBP), com os resíduos de aminoácidos conservados destacados pela cor rosa, além disso, a região que mostra os espaçamentos (*gaps*) na sequência de OsACBP, que se refere a resíduos distintos no alinhamento das sequências. No alinhamento estrutural realizado no PyMol (figura 35) da TcACBP com energia minimizada (KO, J. et al., 2012) e OsACBP, o RMSD (derivação da raiz quadrada média) foi de 0.260 Â, valor que demonstra mínimas divergências entre o molde e o modelo (MADHAVI SASTRY et al., 2013).

```
T-COFFEE, Version_11.00 (Version_11.00)
Cedric Notredame
CPU TIME:0 sec.
SCORE=100
*
 BAD AVG GOOD
*
EOY25439.1        : 100
5H3I_1|Chains     : 100
cons              :  10

TcACBP  EOY25439.1      ITRRWSLRLQFIYCERKMGLKEEFEEYAEKAKTLPETTTNENKLILYGLFKQATVGPVNTSRPGMF
OsACBP  5H3I_1|Chains   --RWGS-----------MGLQEEFEEFAEKAKTLPDTISNEDKLLLYGLYKQATVGPVTTGRPGIF

        cons              *  *            ***:*****:********:* :**:**:****:********.*.***:*


TcACBP  EOY25439.1      NMRERYKWDAWKAVEGKSKEEAMSDYITKVKQLLEEAAASS-
OsACBP  5H3I_1|Chains   NLKDRYKWDAWKAVEGKSKEEAMADYITKVKQLLEEASASTS

        cons            *:::*******************:*************:**:
```

Figura 34: Alinhamento entre as sequências de aminoácidos da TcACBP e OsACBP (PDB 5H3I). Fonte: Predição gerada no T-COFFEE, adaptado.

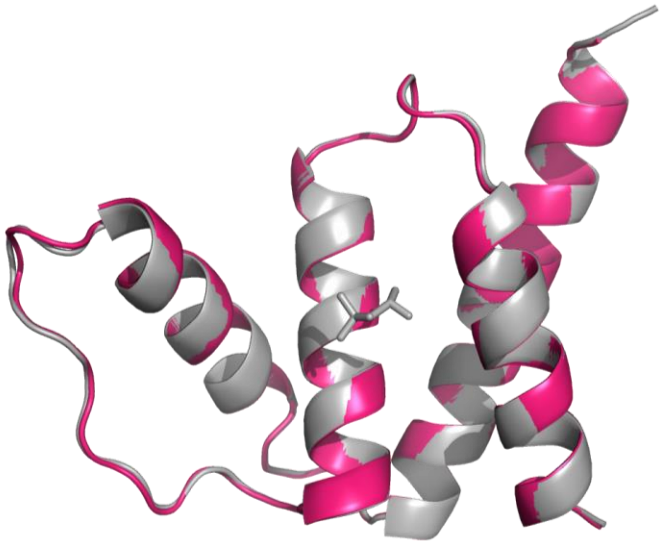

Figura 35: Alinhamento estrutural entre TcACBP – energia minimizada (rosa) e o molde OsACBP PDB ID 5H3I (cinza). Fonte: Imagem gerada pelo programa PyMol; SCHRÖDINGER (2021).

De acordo com a validação do modelo construído, na figura 36 (a) o valor apresentado do QMEAN que fornece estimativas de qualidade absoluta globais (para toda estrutura), e locais (por resíduos) apresentado foi de 0.84; e o GMQE que se refere a estimativa de qualidade do modelo foi de 0.75, demonstrando uma alta qualidade do modelo construído, quando comparado com estruturas resolvidas por métodos experimentais (SCHWEDE et al., 2003; MANDHAVI SASTRY et al., 2013; WATERHOUSE et al., 2018). Em seguida, o modelo foi

validado também utilizando o PROCHEK, figura 36 (b), que através do gráfico de Ramachandran gerado demonstrou que 97,5% dos resíduos estão em regiões energeticamente mais favoráveis, e 2,5% em regiões permitidas, portanto, o modelo é considerado de boa qualidade (acima de 90%) (LASKOWSI, R.A.*et al.*, 1993).

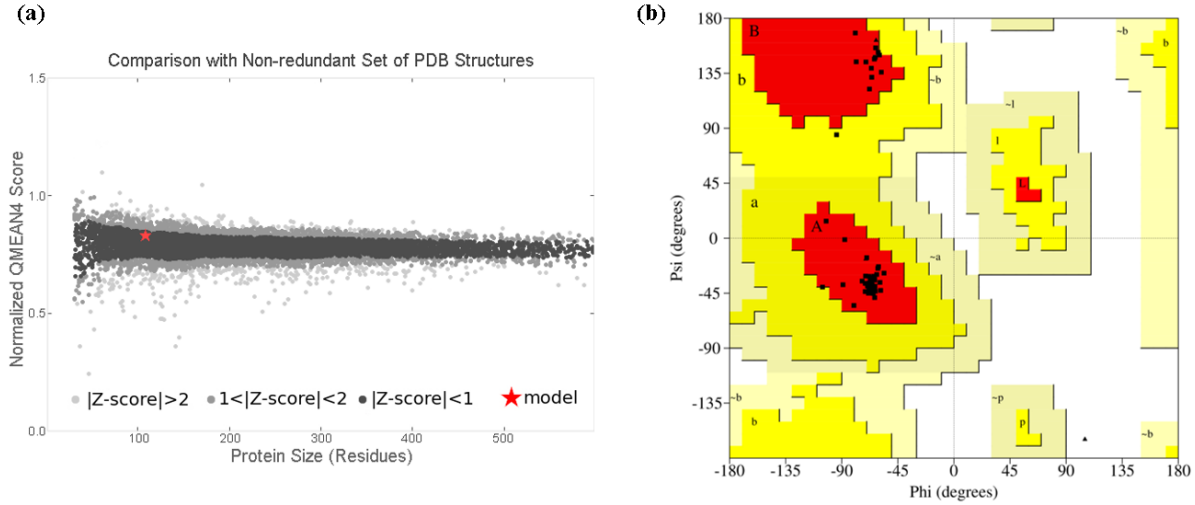

Figura 36: Validação do modelo TcACBP. (a) QMEAN e (b) PROCHEK. **Fonte:** Predição SWISS-MODEL Workspace e PROCHECK.

5.4 Identificação da alça divergente em TcACBP

No alinhamento múltiplo entre as sequências de aminoácidos da TcACBP, MpACBP e MrACBP realizado no T-COFFEE (DI TOMMASO et al., 2011), com o resultado observamos (figura 37) o SCORE (qualidade do alinhamento estre as sequências) de 100% para TcACBP, MpACBP e MrACBP; o "*" indicam as regiões conservadas entre as sequências e uma maior similaridade e conservação dos resíduos de aminoácidos ao ser demonstrado pela cor rosa; e no alinhamento estrutural realizado no PyMOL (SCHRÖDINGER, 2021) (figura 38) mostrou a presença da alça divergente em TcACBP localizado a partir do resíduo de aminoácido 35 ao 40 e com *gap* entre os resíduos (PET----TTN). A região de espaçamentos (*gap* ou *mismatch)* na sequência da TcACBP indicando que nesta região apresenta resíduos de aminoácidos distintos das demais sequências, podendo indicar uma possível mutação pontual de inserções ou deleções introduzidas nesta região. Os "*" indicando que os resíduos são similares entre as sequências.

```
T-COFFEE, Version_11.00 (Version_11.00)
Cedric Notredame
CPU TIME:0 sec.
SCORE=100
*
 BAD AVG GOOD
*
TcACBP  EOY25439.1      : 100
MpACBP  3FP5 1|Chain    : 100
MrACBP  ESK94037.1      : 100
        cons            :  10

TcACBP  EOY25439.1      ITRRWSLRLQFIYCERKMGLKEEFEEYAEKAKTIPET----TTNENKLILYGLFKQATVGPVNTSR
MpACBP  3FP5 1|Chain    -----S-----------HMSKAKFDKAVEIVQSIPKDGPIKPTQDEQLYFYKYFKQATVGDVNISR
MrACBP  ESK94037.1      ------------------MSKAKFDKAVKIVQSIPKDGPIQPTQEEQLYFYKYYKQATVGDVNIPR

        cons                              * :*:: .: .::**:     .*::::* :*  :****** ** .*


TcACBP  EOY25439.1    45 PGMFNMRERYKWDAWKAVEGKSKEEAMSDYITKVKQLLEEAAASS-----------
MpACBP  3FP5 1|Chain  56 PGLMDFTGKAKWDAWKSVEGTSKEVAYQKYVEKLLEILKKADTEESKKYIAEIEAA
MrACBP  ESK94037.1    56 PGLIDFTGKAKWDAWKSVEGTSTEDAYQKYVETLLEILKKADTEESKKYIAEIEAA

        cons            **:::: : ******:***.*.* * ..*: .: ::*::* :..
```

Figura 37: Alinhamento múltiplo entre as sequências de aminoácidos de TcACBP, MpACBP e MrACBP. Fonte: Predição gerada no T-COFFEE, adaptado.

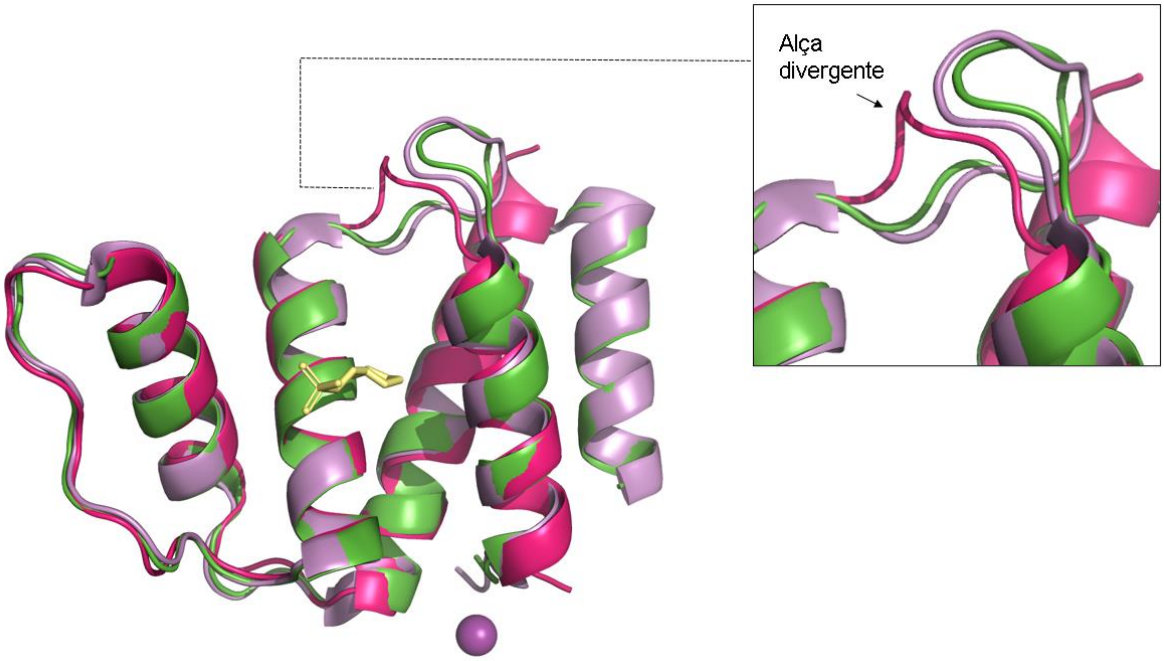

Figura 38: Alinhamento estrutural entre TcACBP (rosa), MpACBP (lilás) e MrACBP (verde). Identificação de alça divergente em TcACBP (rosa), indicado pela seta. Esfera na cor roxa indicando íon zinco (ZN^{2+}). Molécula MES ou ácido 2- (*N*-morfolino) etanossulfônico (amarelo), localizada no sítio de ligação CoA em MpACBP. **Fonte:** Imagem gerada pelo PyMol; SCHRÖDINGER (2021).

Portanto, observamos no alinhamento estrutural realizado no PyMol (SCHRÖDINGER, 2021) (figura 38) que as proteínas homólogas que compartilham a evolução a partir de um ancestral comum, apresentaram estruturalmente regiões internas conservadas, constituídas pela estrutura secundária em α-hélices porém, as principais diferenças estruturais

encontradas ocorreram na região externa, que constituem principalmente por alças (loops) que por sua vez ligam as α-hélices (CALIXTO, 2013; SANTOS FILHO; BICCA DE ALENCASTRO, 2003), onde alça de TcACBP apresentou divergente em relação as alças de MrACBP e MpACBP.

5.5 Triagem virtual baseada no receptor e estudo de *Docking* entre MrACBP e ligantes

Além dos inibidores que foram adquiridos descritos por Fritzler & Zhu (2011), fizemos a busca por drogas conhecidas ou similares no DrugBank que atuam como ligantes e prováveis inibidores da ACBP, a partir da triagem das moléculas encontradas desconsideramos as moléculas que não foram possíveis gerar resultados (em "cinza" na tabela 3); selecionamos as demais 24 moléculas candidatas para prosseguirem com os testes. Com os resultados gerados do acoplamento molecular (*docking*) (tabela 4) realizado com o AutoDock Vina (TROTT; OLSON, 2010), selecionamos 21 moléculas como ligantes positivos ao serem complexados com o receptor (proteína-alvo: MrACBP) na região do sítio ativo e que apresentaram afinidade de ligação com o receptor, desse modo, foram organizadas com os valores de energias de afinidade da melhor para pior.

Tabela 3 - Lista de inibidores contra a proteína de ligação de acil-CoA graxo (ACBP) anti-*Cryptosporidium parvum*, compostos semelhantes, e drogas conhecidas a partir do banco de dados DrugBank. Em destaque (cinza) compostos que não obtiveram resultados.

Continua

Nomes dos compostos	PubChem CID	Compostos semelhantes	PubChem CID
1S, 2R-fenilpropanolamina HCl	6433169		
acetazolamida	1986		
bitionol	2406		
broxiquinolina	2453		
clorpromazina	2726		
cloxacilina sódica	23675743	**Cloxacilina**	6098
cloxiquina	2817		
curcumina	969516		
ácido gambógico	9852185		
Homatropina metilbrometo	10429215	**Metilhomatropina**	5282592
hidralazina HCl	9351	**Hidralazina**	3637
acetato de hidrocortisona	5744		
isoxicam	54677972		
sulfossalicilato de meclociclina	54676538	**Meclociclina**	54676539

Tabela 3 - Lista de inibidores contra a proteína de ligação de acil-CoA graxo (ACBP) anti-*Cryptosporidium parvum*, compostos semelhantes, e drogas conhecidas a partir do banco de dados DrugBank. Em destaque (cinza) compostos que não obtiveram resultados.

			Conclusão
Nomes dos compostos	**PubChem CID**	**Compostos semelhantes**	**PubChem CID**
cloridrato de mitoxantrona	51082	**Mitoxantrona**	4212
oxacilina sódica	23667631		
cloridrato de fenazopiridina	8691		
sulfato de fenelzina	61100		
fenotrina	4767		
fenitoína de sódio	657302		
pregnenolona	8955		
pristimerin	159516		
quinalizarina	5004		
rifampicina	135398735		
rifaximina	6436173		
desidrocolato de sódio	23675007		
estreptozocina	29327		
tirotricina	452550		

Fonte: Adaptado de FRITZLER; ZHU (2011)

Analisamos o complexo formado entre cada um dos ligantes com o receptor, no PyMol 2.5.2 e determinamos qual posição (*pose*) possuía maior número de interações entre ligante e receptor, além de concentrar o ligante mais próximo ao sítio ativo (tabela 4). Em seguida, foram selecionados os cinco melhores e classificados como TOP 5 (figura 39), considerando as melhores energias de afinidade, o maior número de interações entre os aminoácidos do ligante com o receptor, localização quanto a proximidade no sítio ativo; as moléculas nesse grupo foram: Ácido gambogico (-7.3 Kcal/mol), Isoxicam (-7.3 Kcal/mol), Meclociclina (-7.1 Kcal/mol), Cloxacilina (-7.1 Kcal/mol), e Pregnenolona (-6.9 Kcal/mol) (tabela 3). Portanto, selecionamos estes ligantes para prosseguir com as análises, por possuírem maior número de interações e melhor energia de afinidade.

Tabela 4 - Ligantes positivos que realizaram *docking* molecular com a MrACBP (receptor) no AutoDock Vina e que foram complexados com o receptor no PyMol. Os cinco melhores ligantes (TOP 5) em destaque – cinza, foram considerados de acordo com a *pose* (posição) dos ligantes no sítio de acordo com parâmetros de energia de afinidade (gerada no *docking)*, localização quanto a proximidade no sítio ativo e presença do maior número de ligações entre os aminoácidos do ligante-receptor.

Nomes dos compostos	PubChem CID	Posição	Energia de afinidade (Kcal/mol)
Ácido gambogico	**9852185**	**1/9**	**-7.3**
Isoxicam	**54677972**	**1/9**	**-7.3**
Meclociclina	**54676539**	**1/9**	**-7.1**
Cloxacilina	**6098**	**1/9**	**-7.1**
Pregnenolona	**8955**	**1/9**	**-6.9**
Curcumina	969516	1/9	-6.7
Metilhomatropina	5282592	2/9	-6.6
Acetato de hidrocortisona	5744	3/9	-6.6
Quinalizarina	5004	1/9	-6.5
Fenotrina	4767	7/9	-6.4
Pristimerina	159516	9/9	-6.2
Rifampicina	135398735	1/1	-6.0
Bitionol	2406	3/9	-5.9
Hidralazina	3637	1/9	-5.8
Broxiquinolina	2453	1/9	-5.6
Clorpromazina	2726	1/9	-5.6
Mitoxantrona	4212	1/9	-5.6
Cloxiquina	2817	1/9	-5.5
Rifaximina	6436173	1/2	-5.5
Estreptozocina	29327	1/9	-5.2
Acetazolamida	1986	2/9	-4.6

Fonte: A autora.

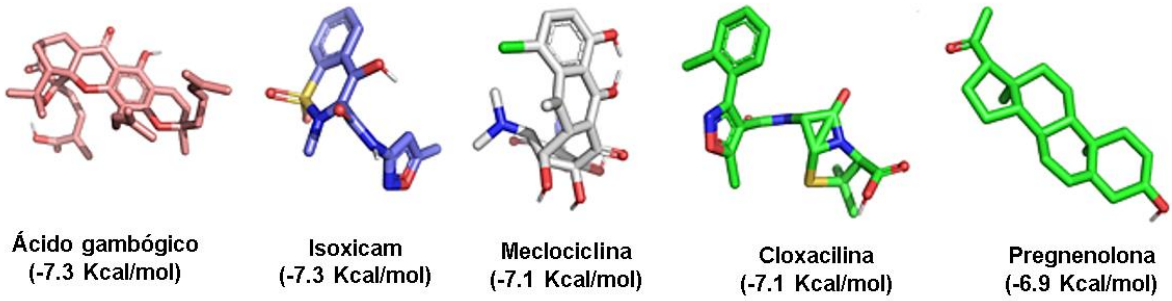

Figura 39: Estruturas 3D do TOP 5 inibidores selecionados com maior energia de afinidade. Na ordem da esquerda para direita do mais eletronegativo para o menos eletronegativo: o ácido gambogico, isoxicam, meclociclina, cloxacilina e pregnenolona. **Fonte:** Predição das estruturas 3D pelo Marvin Sketch v. 19.25.

No estudo do complexo formado entre MrACBP (proteína) e o Ácido gambógico (ligante CID 9852185), mostrando o ligante localizado na região do sítio ativo com melhor *pose* gerada 1/9 realizado no PyMol 2.5.2 (SCHRÖDINGER, 2021) (figura 40 A), e com energia de afinidade de -7.3 Kcal/mol no resultado do *docking*. Observamos no mapa de interação 3D

(figura 40 B) gerado pelo PLP Web Service (ADASME et al., 2021), a presença de interações hidrofóbicas que ocorre entre os aminoácidos do sítio GLN 14, TYR 33, TYR 36 e o ligante, e da ligação de hidrogênio entre o aminoácido do sítio GLN 14 e o ligante. As demais interações hidrofóbicas entre os aminoácidos LYS 18, GLN 29, PHE 32 com o ligante e a ligação de hidrogênio com GLN 29 e o ligante se estabilizam ao redor do sítio.

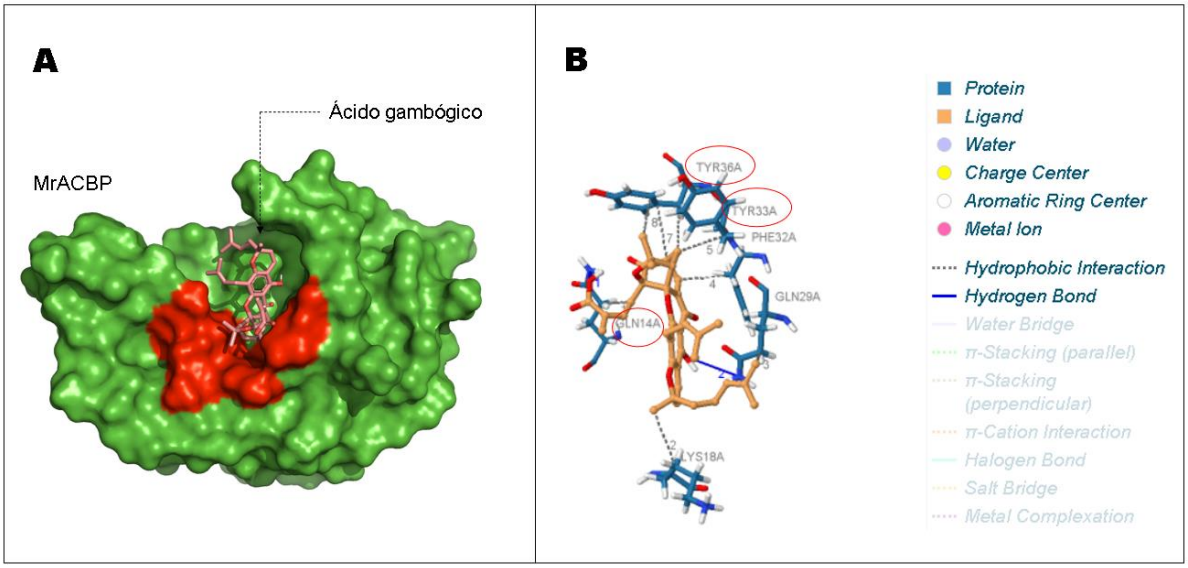

Figura 40: (A) **Complexo formado pela proteína ACBP de *M. roreri* (verde) e o ácido gambogico (ligante CID 9852185).** Mostrando o ligante (rosa) com a *pose* localizado na região do sítio ativo (vermelho). (B) Mapa de interação 3D da proteína (azul) e o ácido gambogico (laranja) com ligações de hidrogênio e interações hidrofóbicas entre eles. **Fonte:** (A) Imagem gerada pelo PyMol; SCHRÖDINGER (2021); (B) Predição PLP Web Service ADASME et al. (2021), adaptado.

O complexo formado entre MrACBP (proteína) e o Isoxicam (ligante CID 54677972), mostrando o ligante localizado na região do sítio ativo com melhor *pose* gerada 1/9 realizado no PyMol 2.5.2 (SCHRÖDINGER, 2021) (figura 41 A), e com energia de afinidade de -7.3 Kcal/mol no resultado do *docking*. Observamos no mapa de interação 3D (figura 41 B) gerado pelo PLP Web Service (ADASME et al., 2021), a presença de interações hidrofóbicas entre o aminoácido do sítio TYR 33, TYR 36 e o ligante, duas ligações de hidrogênio entre os aminoácidos do sítio GLN 14, TYR 78 e o ligante. A presença de uma ligação do tipo Pi-Pi Stacking (ou interações de empilhamento $\pi - \pi$) que ocorre na disposição face-a-face do anel aromático do aminoácido do sítio TYR 36 e do anel aromático do ligante é considerada energeticamente desfavorável, pelo modelo proposto por Hunter & Sander (1990), devido à forte repulsão entre as nuvens de elétrons Pi. Esse empilhamento Pi, é um caso raro, e é possível entre aromáticos ricos em elétrons ou deficientes. Essa interação de empilhamento ajuda na

ligação de ligantes hidrofóbicos ao sítio ativo de enzimas (MOLCANOV & KOJIC-PRODIC, 2019). Além disso, uma interação hidrofóbica entre GLN 29 e o ligante fora da região do sítio ativo.

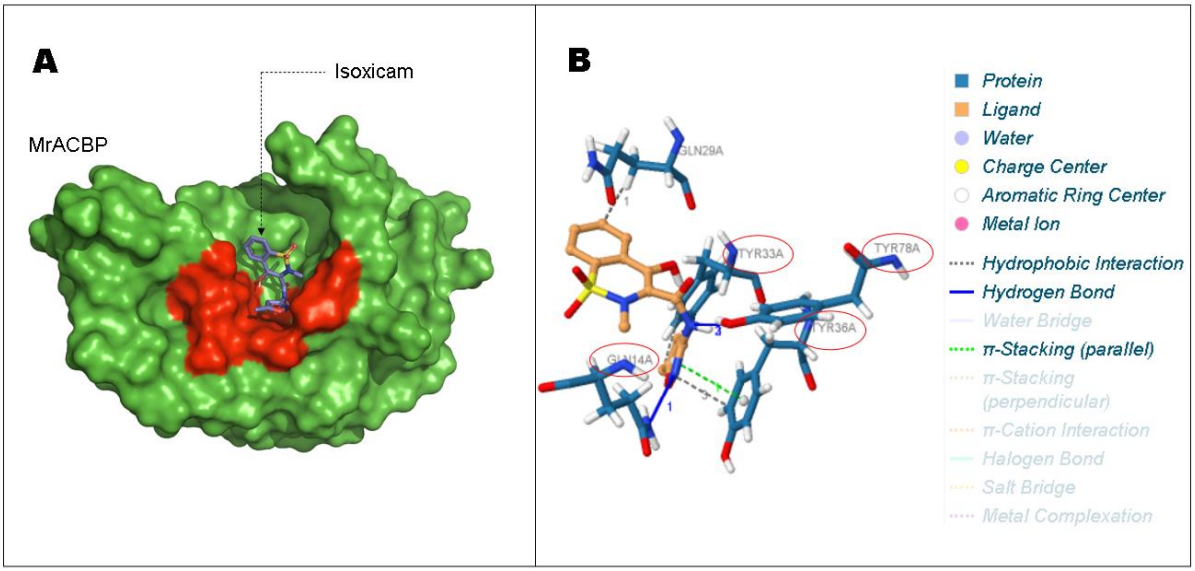

Figura 41: (A) **Complexo formado pela proteína ACBP de *M. roreri* (verde) e o Isoxicam (ligante CID 54677972).** Mostrando o ligante (roxo) com a *pose* localizado na região do sítio ativo (vermelho). (B) Mapa de interação 3D da proteína (azul) e o Isoxicam (laranja) com ligações de hidrogênio, interações hidrofóbicas e uma ligação Pi-Pi Stacking (traço pontilhado verde), entre eles. **Fonte:** (A) Imagem gerada pelo PyMol, SCHRÖDINGER (2021); (B) Predição PLP Web Service, ADASME et al., (2021), adaptado.

O complexo formado entre MrACBP (proteína) e o Meclociclina (ligante CID 54676539), mostrando o ligante com a *pose* localizado na região do sítio ativo com melhor *pose* gerada 1/9 realizado no PyMol 2.5.2 (SCHRÖDINGER, 2021) (figura 42 A), e com energia de afinidade de -7.1 Kcal/mol no resultado do *docking*. Observamos no mapa de interação 3D (figura 42 B) gerado pelo PLP Web Service (ADASME et al., 2021), a presença de interações hidrofóbicas entre os aminoácidos do sítio TYR 33, TYR 36 e o ligante, e ligações de hidrogênio entre os aminoácidos do sítio TYR 33, LYS 59, TYR 78 e o ligante. A ligação de hidrogênio entre GLN 29 e o ligante se estabiliza ao redor do sítio.

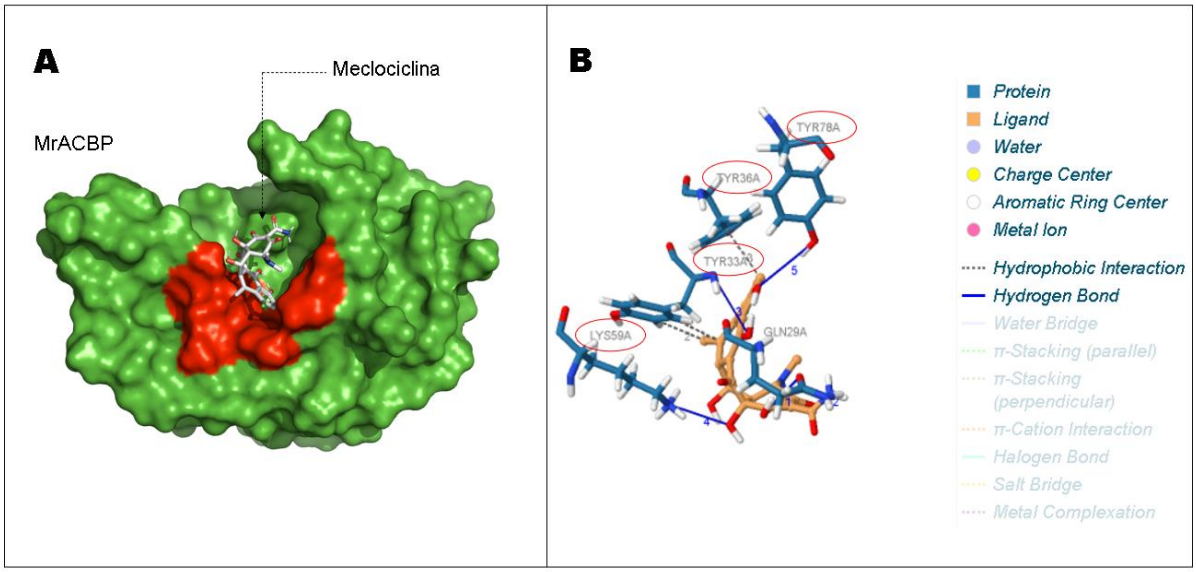

Figura 42: (A) **Complexo formado pela proteína ACBP de *M. roreri* (verde) e o Meclociclina (ligante CID 54676539).** Mostrando o ligante (cinza) com a *pose* localizado na região do sítio ativo (vermelho). (B) Mapa de interação 3D da proteína (azul) e a meclociclina (laranja) com ligações de hidrogênio e interações hidrofóbicas entre eles. **Fonte:** (A) Imagem gerada pelo PyMol; SCHRÖDINGER (2021); (B) Predição PLP Web Service, ADASME et al., (2021), adaptado.

O complexo formado entre MrACBP (proteína) e o Cloxacilina (ligante CID 6098), mostrando o ligante com a *pose* localizado na região do sítio ativo com melhor *pose* gerada 1/9 realizado no PyMol 2.5.2 (SCHRÖDINGER, 2021) (figura 43 A), e com energia de afinidade de -7.1 Kcal/mol no resultado do *docking*. Observamos no mapa de interação 3D (figura 43 B) gerado pelo PLP Web Service (ADASME et al., 2021), a presença de ligações de hidrogênio entre os aminoácidos do sítio GLN 14, TYR 78 e o ligante. As demais, interação hidrofóbica entre LEU 16 e o ligante, e a ligação de hidrogênio entre GLN 29 e o ligante se estabilizam ao redor do sítio.

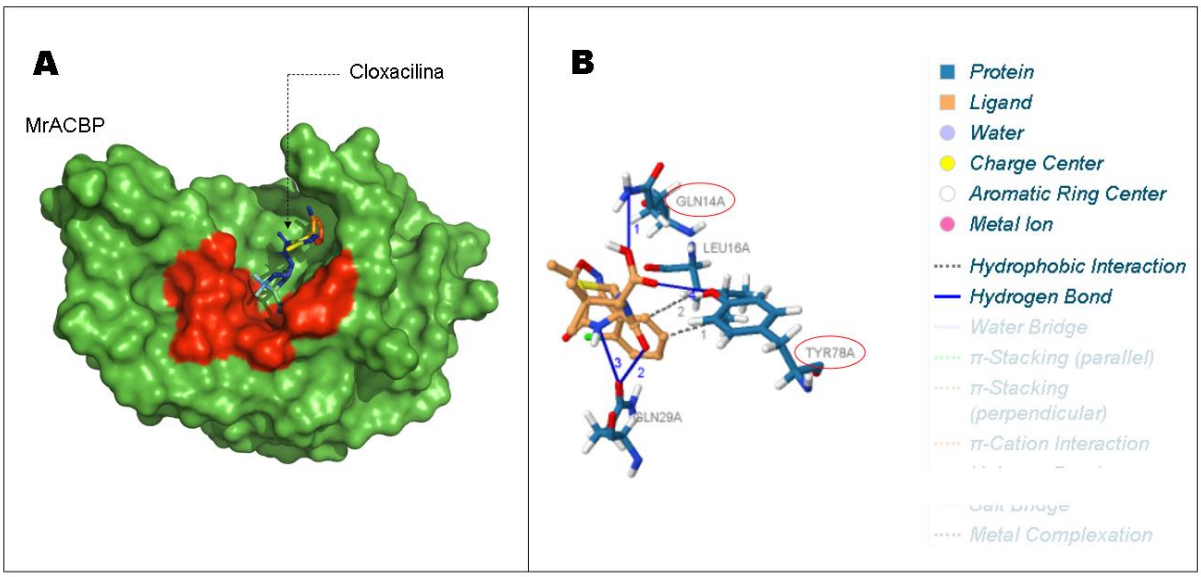

Figura 43: (A) **Complexo formado pela proteína ACBP de *M. roreri* (verde) e o Cloxacilina (ligante CID 6098).** Mostrando o ligante (verde) com a *pose* localizado na região do sítio ativo (vermelho). (B) Mapa de interação 3D da proteína (azul) e o cloxacilina (laranja) com ligações de hidrogênio e interações hidrofóbicas entre eles. **Fonte:** (A) Imagem gerada pelo PyMol; SCHRÖDINGER (2021); (B) Predição PLP Web Service, ADASME et al. (2021), adaptado.

O complexo formado entre MrACBP (proteína) e o Pregnenolona (ligante CID 8955), mostrando o ligante com a *pose* localizado na região do sítio ativo com melhor *pose* gerada 1/9 realizado no PyMol 2.5.2 (SCHRÖDINGER, 2021) (figura 44 A), e com energia de afinidade de -6.9 Kcal/mol no resultado do *docking*. Observamos no mapa de interação 3D (figura 44 B) gerado pelo PLP Web Service (ADASME et al., 2021), a presença de interações hidrofóbicas entre os aminoácidos do sítio GLN 14, TYR 33, TYR 36 e o ligante, ligação de hidrogênio entre o aminoácido do sítio GLN 14 e o ligante. As demais interações hidrofóbicas entre GLN 26, GLN 29 e o ligante ocorreram ao redor do sítio.

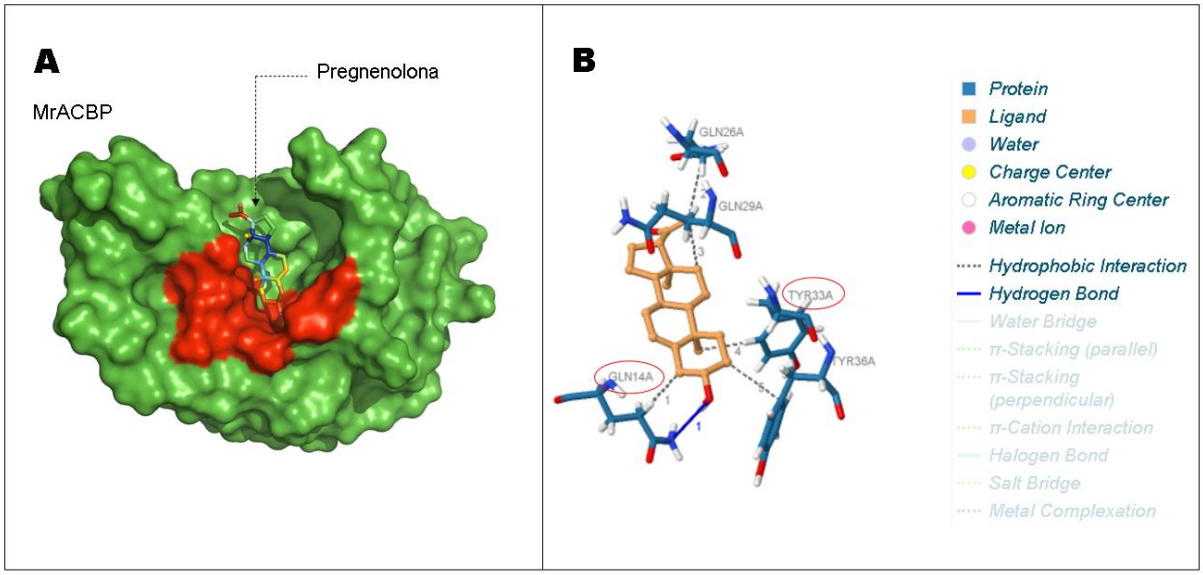

Figura 44: (A) **Complexo formado pela proteína ACBP de *M. roreri* (verde) e o Pregnenolona (ligante CID 8955).** Mostrando o ligante (verde) com a *pose* localizado na região do sítio ativo (vermelho). (B) Mapa de interação 3D da proteína (azul) e o pregnenolona (laranja) com ligações de hidrogênio e interações hidrofóbicas entre eles. **Fonte:** (A) Imagem gerada pelo PyMol; SCHRÖDINGER (2021); (B) Predição PLP Web Service, ADASME et al. (2021), adaptado.

5.6 Estudo de *Docking* entre TcACBP e ligantes

A partir dos resultados gerados do acoplamento molecular (*docking*) realizado com o AutoDock Vina (TROTT; OLSON, 2010) (tabela 4) das 21 moléculas que apresentaram como ligantes positivos, complexamos com o receptor TcACBP para observarmos o nível em que essas moléculas interagem, e organizamos os valores de energia de afinidade da melhor para pior. O sítio ativo da TcACBP foi determinado a partir do alinhamento múltiplo com as sequências de aminoácidos do sítio de MrACBP e MpACBP, descrito por Monzani et al., 2010, realizado no T-COFFEE (DI TOMMASO et al., 2011) e estrutural realizado no PyMol (SCHRÖDINGER, 2021), portanto, os aminoácidos que compõem o sítio ativo de TcACBP são: LYS 32, TYR 47, PHE 50, LYS 51, LYS 73, TYR 92. Além disso, observamos que não houve similaridades entres os aminoácidos do sítio de TcACBP e MrACBP.

Analisamos o complexo formado entre cada um dos ligantes com o receptor TcACBP, no PyMol 2.5.2 (tabela 5) e determinamos qual posição (*pose*) possuía maior número de interações entre ligante e receptor, além de concentrar o ligante mais próximo ao sítio ativo de TcACBP. Desse modo, foram selecionados os três melhores (em cinza, tabela 5) considerando as energias de afinidade, o maior número de interações entre os aminoácidos do ligante com o receptor, localização quanto a proximidade no sítio ativo, as moléculas nesse grupo foram:

Pregnenolona (-7.6 Kcal/mol), Cloxacilina (-7.4 Kcal/mol), Ácido gambógico (-7.3 Kcal/mol). Porém, os ligantes Rifaximina (-8.4 Kcal/mol), e Meclociclina (-8.3 Kcal/mol) não foram considerados devido a maior energia de afinidade que apresentaram na interação com os aminoácidos do sítio ativo que os demais ligantes (tabela 5).

Tabela 5: Ligantes positivos que realizaram *docking* molecular com a TcACBP (receptor) no AutoDock Vina e que foram complexados com o receptor no PyMol. Os três ligantes em destaque (cinza), consideramos de acordo com a posição dos ligantes no sítio ativo de acordo com parâmetros de energia de afinidade (*docking*) para prosseguirem com as análises de dinâmica molecular. Porém, os dois ligantes Rifaximina e Meclociclina com maior energia de afinidade com o sítio de TcACBP de -8.4 Kcal/mol e -8.3 Kcal/mol respectivamente, não foram considerados.

Nomes dos compostos	PubChem CID	Posição	Energia de afinidade (Kcal/mol)
Rifaximina	**6436173**	**1/3**	**-8.4**
Meclociclina	**54676539**	**1/9**	**-8.3**
Pregnenolona	**8955**	**1/9**	**-7.6**
Cloxacilina	**6098**	**1/9**	**-7.4**
Ácido gambogico	**9852185**	**1/9**	**-7.3**
Quinalizarina	5004	1/9	-7.2
Isoxicam	54677972	1/9	-7.1
Acetato de hidrocortisona	5744	3/9	-7.0
Pristimerina	159516	3/9	-6.8
Fenotrina	4767	2/9	-6.8
Bitionol	2406	1/9	-6.8
Curcumina	969516	6/9	-6.7
Metilhomatropina	5282592	3/9	-6.7
Clorpromazina	2726	8/9	-6.1
Broxiquinolina	2453	1/9	-6.0
Cloxiquina	2817	1/9	-5.9
Mitoxantrona	4212	3/9	-5.8
Hidralazina	3637	1/9	-5.5
Estreptozocina	29327	1/9	-5.4
Acetazolamida	1986	1/9	-5.0
Rifampicina	135398735	6/8	-4.7

Fonte: A autora.

No estudo do complexo formado entre TcACBP (proteína) e o Pregnenolona (ligante CID 8955), mostrando o ligante localizado na região do sítio ativo (com cavidade hidrofóbica) com melhor *pose* gerada 1/9 realizado no PyMol 2.5.2 (SCHRÖDINGER, 2021) (figura 45 A), e com energia de afinidade de -7.6 Kcal/mol no resultado do *docking*. Observamos no mapa de interação 3D (figura 45 B) gerado pelo PLP Web Service, a presença de duas interações

hidrofóbicas com aminoácidos do sítio tirosina (TYR) 47A em duas posições diferentes, e fenilalanina (PHE) 50A em duas posições diferentes interagindo com o ligante e; com as demais interações hidrofóbicas com lisina (LYS) 43A, leucina (LEU) 46A interagindo com o ligante ao redor do sítio; e uma ligação de hidrogênio entre leucina (LEU) 34A e o ligante que estabilizam ao redor do sítio.

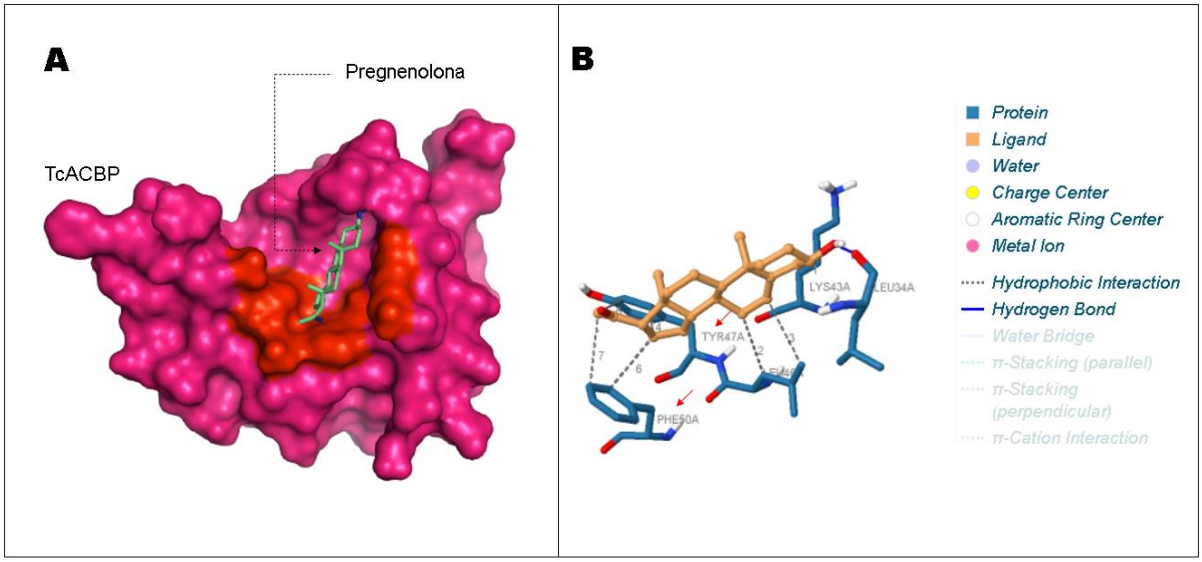

Figura 45: (A) **Complexo formado pela proteína TcACBP (rosa) e o Pregnenolona (ligante CID 8955).** Mostrando o ligante (verde) com a *pose* 1/9 localizado na região do sítio ativo (vermelho). (B) Mapa de interação 3D da proteína (azul) e o pregnenolona (laranja) com ligações de hidrogênio e interações hidrofóbicas entre eles. **Fonte:** (A) Imagem gerada pelo PyMol; SCHRÖDINGER (2021); (B) Predição PLP Web Service, adaptado.

O complexo formado entre TcACBP (proteína) e a Cloxacilina (ligante CID 6098), mostrando o ligante localizado na região do sítio ativo com melhor *pose* gerada 1/9 realizado no PyMol 2.5.2 (SCHRÖDINGER, 2021) (figura 46 A), e com energia de afinidade de -7.4 Kcal/mol no resultado do *docking*. Observamos no mapa de interação 3D (figura 46 B) gerado pelo PLP Web Service (ADASME et al., 2021), a presença de duas ligações de hidrogênio, uma que ocorre entre o aminoácido tirosina (TYR)92A localizado no sítio e o ligante e a segunda fora da região do sítio entre asparagina (ASN) 40A e o ligante; interações hidrofóbicas entre os aminoácidos do sítio tirosina 47A interagem com o ligante em duas posições diferentes; e outras interações hidrofóbicas que estabilizam ao redor do sítio com aminoácidos alanina (ALA) 31A, leucina (LEU) 46A e o ligante. E a presença de uma ligação do tipo Pi-Pi Stacking (ou interações de empilhamento $\pi - \pi$) que ocorre disposta paralelamente e deslocado entre o anel aromático do aminoácido do sítio fenilalanina (PHE) 50A e o anel aromático do ligante, é considerada energeticamente favorável, pelo modelo proposto por Hunter & Sander (1990).

Esse tipo de interação de empilhamento Pi desempenha funções no reconhecimento molecular, contribui com empacotamento cristalino de compostos aromáticos, porém, podem ser moduladas por estímulos externos, e também por modificações químicas (MOLCANOV & KOJIC-PRODIC, 2019).

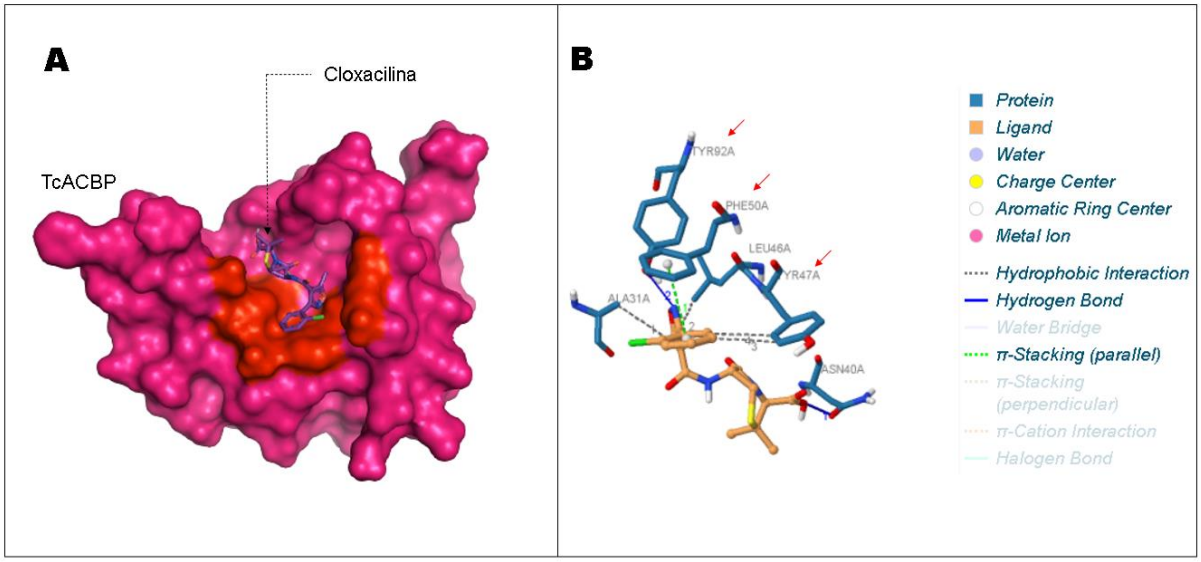

Figura 46: (A) **Complexo formado pela proteína TcACBP (rosa) e o Cloxacilina (ligante CID 6098).** Mostrando o ligante (roxo) com a *pose* 1/9 localizado na região do sítio ativo (vermelho). (B) Mapa de interação 3D da proteína (azul) e o Cloxacilina (laranja) com ligações de hidrogênio, interações hidrofóbicas entre a proteína e o ligante e uma ligação Pi - Stacking entre o anel aromático da proteína e o anel aromático do ligante. **Fonte:** (A) Imagem gerada pelo PyMol; SCHRÖDINGER (2021); (B) Predição PLP Web Service, ADASME et al. (2021), adaptado.

O complexo formado entre TcACBP (proteína) e o Ácido gambógico (ligante CID 9852185), mostrando o ligante localizado na região do sítio ativo com melhor *pose* gerada 1/9 realizado no PyMol 2.5.2 (SCHRÖDINGER, 2021) (figura 47 A), e com energia de afinidade de -7.3 Kcal/mol no resultado do *docking*. Observamos no mapa de interação 3D (figura 47 B) gerado pelo PLP Web Service (ADASME et al., 2021), a presença de três ligações de hidrogênio, uma que ocorre entre o aminoácido lisina (LYS) 73A localizado no sítio e o ligante e as outras duas ligações de hidrogênio com os aminoácidos glutamina (GLN) 36A, lisina (LYS) 43A interagem com o ligante fora da região do sítio; interações hidrofóbicas entre os aminoácidos do sítio tirosina (TYR) 47A, fenilalanina (PHE) 50A interagindo com o ligante, e outras interações hidrofóbicas que estabilizam ao redor do sítio com os aminoácidos alanina (ALA) 31A, lisina (LYS) 43A, leucina (LEU) na posição 44A e 46A e o ligante; além disso, possui uma ponte salina entre o aminoácido lisina (LYS) 32A e o ligante, na região do sítio ativo, indicando a combinação de uma ligação de hidrogênio e a ligação iônica culminando em

duas interações não covalentes, apesar de serem consideradas interações fracas, pequenas interações não covalentes podem contribuir como estabilizadoras da conformação dobrada entropicamente de proteínas.

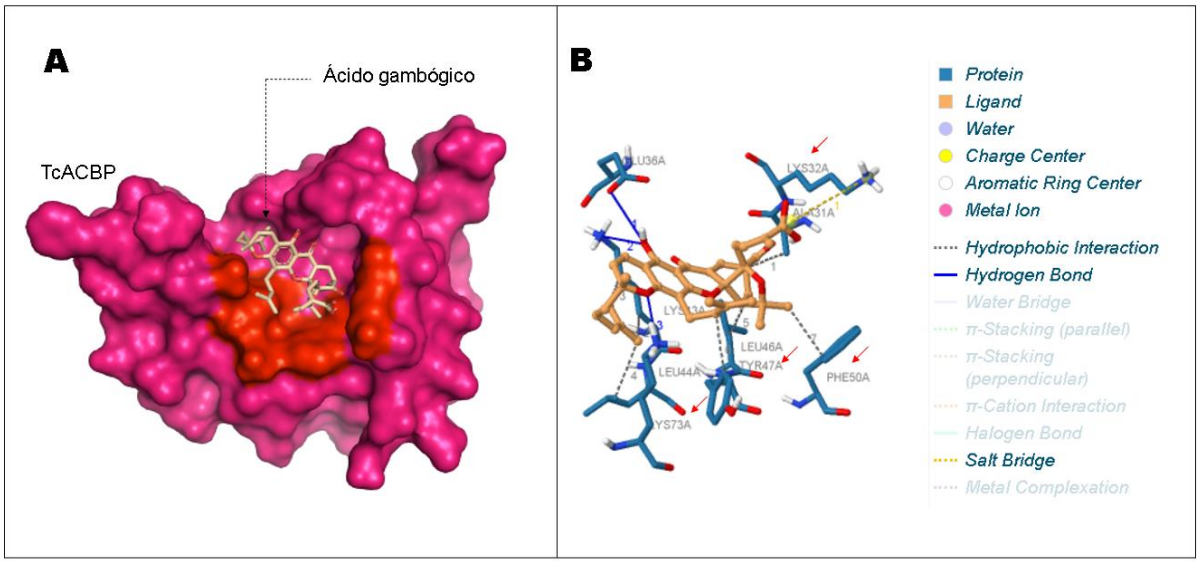

Figura 47: (A) **Complexo formado pela proteína TcACBP (rosa) e o Ácido gambógico (ligante CID 9852185).** Mostrando o ligante (bege) com a *pose* 1/9 localizado na região do sítio ativo (vermelho). (B) Mapa de interação 3D da proteína (azul) e o Ácido gambógico (laranja) com ligações de hidrogênio, interações hidrofóbicas entre a proteína e o ligante e uma ponte salina formada pela interação não covalente entre o aminoácido lisina (LYS) 32A da proteína e o ligante. **Fonte:** (A) Imagem gerada pelo PyMol; SCHRÖDINGER (2021); (B) Predição PLP Web Service, ADASME et al. (2021), adaptado.

Ao comparamos a afinidade entre MrACBP e os cinco ligantes de maior interação com os aminoácidos do sítio ativo e suas energias de afinidade com a interação destes ligantes e o TcACBP (tabela 6), observamos que os ligantes Meclociclina, Pregnenolona e Cloxacilina de acordo com suas energias de afinidade de -8.3 Kcal/mol,-7.6 Kcal/mol e -7.4 Kcal/mol, respectivamente do resultado de *docking* obtiveram maior afinidade com TcACBP do que com MrACBP, o ligante Ácido gambógico com -7.3 Kcal/mol teve o mesmo valor de afinidade com MrACBP e TcACBP. Porém, o Isoxicam apresentou maior afinidade com MrACBP com -7.3 Kcal/mol do que com TcACBP que teve -7.1 Kcal/mol.

Tabela 6 – Comparação da afinidade entre os complexos de MrACBP e os cinco ligantes de maior interação e suas energias de afinidade, com TcACBP e os mesmos ligantes. Listados do ligante de maior energia de afinidade (mais eletronegativo) para o de menor afinidade (menos eletronegativo) dos resultados do *docking*.

Continua

Complexos	Energia de afinidade	Complexos	Energia de afinidade
MrACBP[a] e Ácido gambógico[c]	-7.3 Kcal/mol[d]	TcACBP[b] e Meclociclina[c]	-8.3 Kcal/mol[d]

Tabela 6 – Comparação da afinidade entre os complexos de MrACBP e os cinco ligantes de maior interação e suas energias de afinidade, com TcACBP e os mesmos ligantes. Listados do ligante de maior energia de afinidade (mais eletronegativo) para o de menor afinidade (menos eletronegativo) dos resultados do *docking*.

Conclusão

Complexos	**Energia de afinidade**	**Complexos**	**Energia de afinidade**
MrACBP[a] e Isoxicam[c]	-7.3 Kcal/mol[d]	TcACBP[b] e Pregnenolona[c]	-7.6 Kcal/mol[d]
MrACBP[a] e Meclociclina[c]	-7.1 Kcal/mol[d]	TcACBP[b] e Cloxacilina[c]	-7.4 Kcal/mol[d]
MrACBP[a] e Cloxacilica[c]	-7.1 Kcal/mol[d]	TcACBP[b] e Ácido gambógico[c]	-7.3 Kcal/mol[d]
MrACBP[a] e Pregnenologa[c]	-6.9 Kcal/mol[d]	TcACBP[b] e Isoxicam[c]	-7.1 Kcal/mol[d]

Fonte: A autora
[a]Proteína de ligação a acil CoA (ACBP) de *Moniliophthora roreri*
[b] Proteína de ligação a acil CoA (ACBP) de *Theobroma cacao*
[c]Nomes dos compostos (ligantes)
[d]energia de afinidade de interação do resultado de *docking*

Realizamos o estudo do complexo formado entre TcACBP (proteína) e o Isoxicam (ligante CID54677972), mostrando o ligante localizado na região do sítio ativo com melhor *pose* gerada 1/9 realizado no PyMol 2.5.2 (SCHRÖDINGER, 2021) (figura 48 A), e com energia de afinidade de -7.1 Kcal/mol no resultado do *docking*. Observamos no mapa de interação 3D (figura 48 B) gerado pelo PLP Web Service (ADASME et al., 2021), a presença de duas ligações de hidrogênio, uma que ocorre entre o aminoácido TYR (tirosina) 92 localizado no sítio e o ligante e a segunda fora da região do sítio entre LYS (lisina) 43 e o ligante; interações hidrofóbicas entre os aminoácidos do sítio PHE (fenilalanina) 50 interagem com o ligante; e outra interação hidrofóbica que estabilizam ao redor do sítio com aminoácidos LYS (lisina) 43 e o ligante. E a presença de uma ligação do tipo Pi-Pi Stacking (ou interações de empilhamento $\pi - \pi$) que ocorre disposta paralelamente e deslocado entre o anel aromático do aminoácido do sítio PHE (fenilalanina) 50 e o anel aromático do ligante, é considerada energeticamente favorável, pelo modelo proposto por Hunter & Sander (1990). Esse tipo de interação de empilhamento Pi desempenha funções no reconhecimento molecular, contribui com empacotamento cristalino de compostos aromáticos, porém, podem ser moduladas por estímulos externos, e também por modificações químicas (MOLCANOV & KOJIC-PRODIC, 2019).

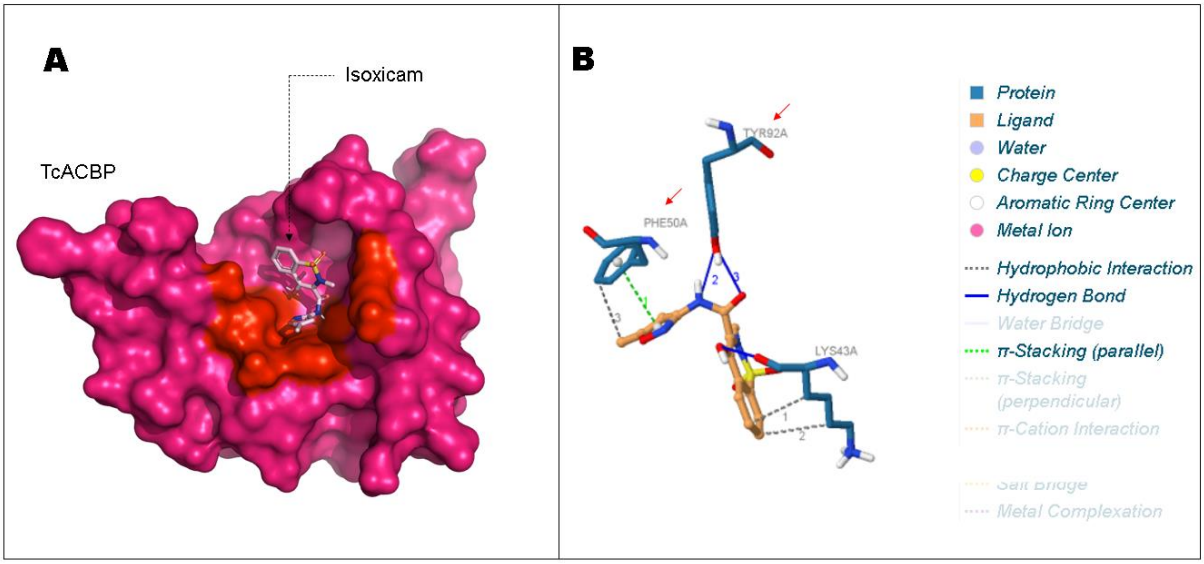

Figura 48: (A) **Complexo formado pela proteína TcACBP (rosa) e o Isoxicam (ligante CID 54677972).** Mostrando o ligante com a *pose* 1/9 localizado na região do sítio ativo (vermelho). (B) Mapa de interação 3D da proteína (azul) e o isoxicam (laranja) com ligações de hidrogênio, interações hidrofóbicas entre a proteína e o ligante e uma ligação Pi - Stacking entre o anel aromático da proteína e o anel aromático do ligante. **Fonte:** (A) Imagem gerada pelo PyMol; SCHRÖDINGER (2021); (B) Predição PLP Web Service, ADASME et al. (2021), adaptado.

Portanto, esses resultados sugerem que os ligantes e prováveis inibidores não se mostraram seletivos na especificidade de interação com MrACBP. Apesar de buscarmos inicialmente moléculas com potencial de afinidade de ligação e inibição seletivo a MrACBP, o papel dos potenciais efeitos fisiológicos destes inibidores em MrACBP e TcACBP precisam ser mais explorados por meio de testes *in vitro* e *in vivo*.

5.7 Dinâmica Molecular de estabilidade entre MrACBP e Isoxicam

Realizamos o estudo da dinâmica molecular de estabilidade do complexo formado pela proteína ACBP de *M. roreri* e o ligante Isoxicam. Foi gerado o gráfico RMSD (desvio da raiz quadrática média, do inglês *root mean square deviation*), figura 49, da cadeia principal, envolvendo a trajetória da simulação em função do tempo (ps). O RMSD calculou a cadeia principal (ou *backdone*) da proteína MrACBP com o ligante Isoxicam, sem os átomos de hidrogênio, permanecendo os átomos pesados (carbono, nitrogênio, oxigênio). Portanto, avaliou o encaixe do ligante no sítio ativo em relação à cadeia principal da proteína, o ideal é não desviar mais de 0,1 nanometro até o final da dinâmica molecular, porém o desvio ocorreu mais de 0,1 nanometro o que significa que desviou muito o ligante da proteína, resultando pra esse inibidor o RMSD foi considerado muito alto, que pode indicar que o ligante esteve saindo

do sítio ativo, além disso, mostrou uma variação muito grande, e em vários pontos do meio da simulação não conseguiu estabilizar.

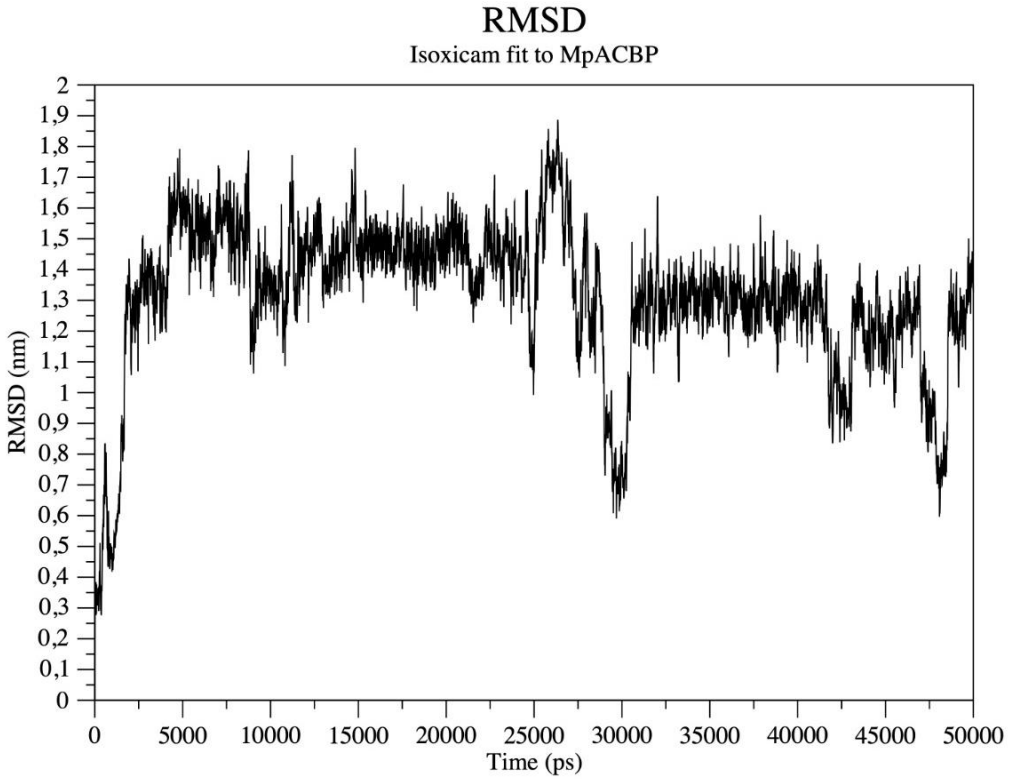

Figura 49: Gráfico RMSD (derivação da raiz quadrática média) da dinâmica de estabilidade entre MrACBP e o ligante Isoxicam. **Fonte:** Predição GROMACS 2020.

No cálculo da energia das interações eletrostáticas (interações não ligadas) entre os átomos do ligante (Isoxicam) e todos os átomos do sistema (receptor – MrACBP) o gráfico (figura 50) demonstrou que nos primeiros 11000 ps (picossegundos) a energia obtida pelo ligante ultrapassava valores acima de 0 KJ/mol demonstrando pouca afinidade com a proteína, porém depois de aproximadamente 11000 picossegundos apresentou estabilidade com a média foi de -28.7241 até o final da simulação, de modo a obter menores energias eletrostáticas, demonstrando que com um tempo mais longo de simulação poderia confirmar uma maior afinidade do ligante ao sítio ativo.

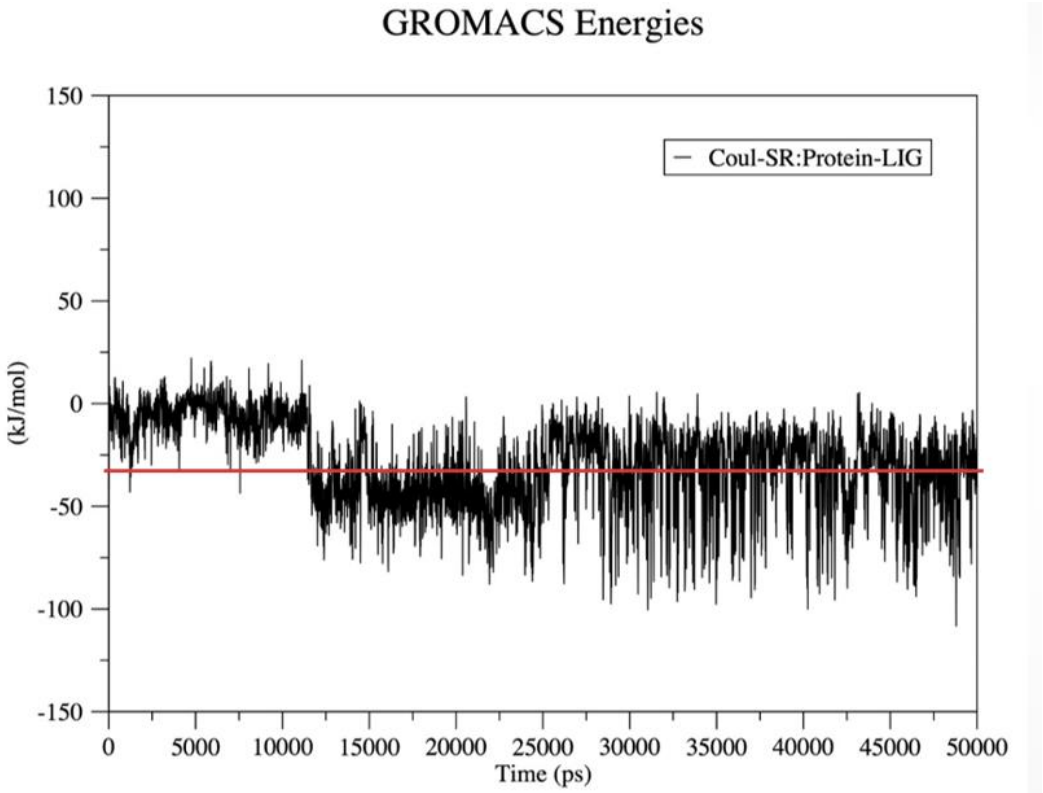

Figura 50: Gráfico do Potencial de Energia (Coulomb) da dinâmica de estabilidade entre MrACBP e o ligante Isoxicam. **Fonte:** Predição GROMACS 2020.

No cálculo da energia das interações de van der Waals (interações não ligadas) entre os átomos do ligante (Isoxicam) e todos os átomos do sistema (receptor – MrACBP) o gráfico (figura 51) demonstrou que a média foi de -72.9937 KJ/mol demonstrando estabilidade e afinidade com a proteína MrACBP, uma vez que a energia obtida pelo ligante não ultrapassava valores acima de 0 KJ/mol.

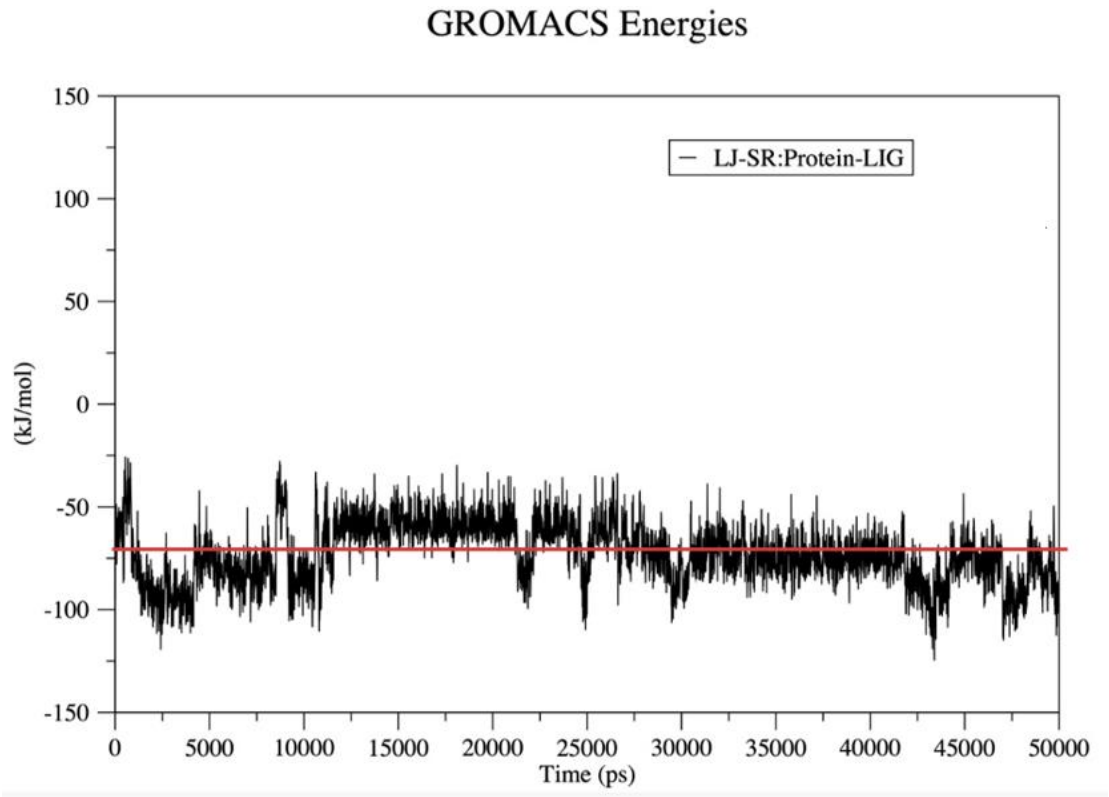

Figura 51: Gráfico do Potencial de Energia (Lennard-Jones) da dinâmica de estabilidade entre MrACBP e o ligante Isoxicam. **Fonte:** Predição GROMACS 2020.

No gráfico de estabilidade das ligações de hidrogênio (figura 52), mostrou que em todo tempo da simulação da proteína MrACBP realizou ligações de hidrogênio com o ligante Isoxicam. Porém não quer dizer que foi somente com aminoácidos da região do sítio, pode ter interagido com ligações de hidrogênio com outros aminoácidos fora da região do sítio também. Nos picos desse gráfico nos tempos de 21000 ps, 25000 ps, 26000 ps e 41000ps, realizou o maior número de ligações de hidrogênio, foram 3 ligações de hidrogênio em cada um desses tempos, isso é favorável uma vez que mantém mais estável a interação da proteína com o ligante. E a quantidade média de 1 ligação de hidrogênio ficou estável até o final da simulação.

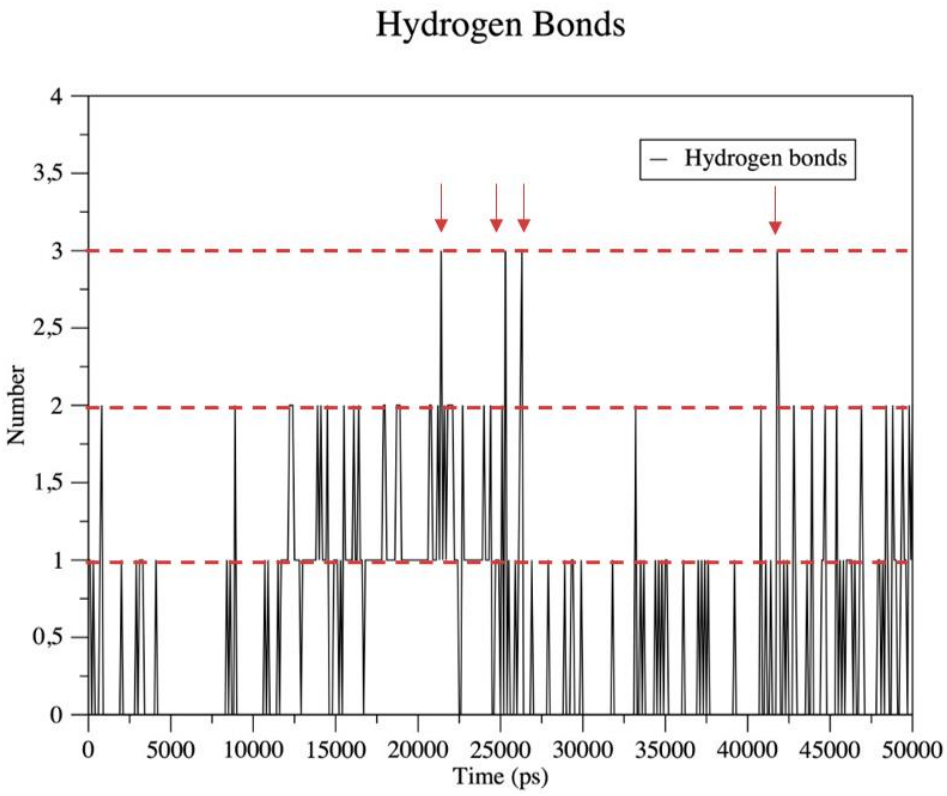

Figura 52: Gráfico da dinâmica molecular de estabilidade entre MrACBP e o ligante Isoxicam – Ligação de hidrogênio. **Fonte:** Predição GROMACS 2020.

No gráfico RMS (desvio da raiz quadrática média, do inglês *root mean square deviation*) de flutuação (figura 53) da simulação de dinâmica molecular, apresentou nas regiões do sítio ativo de MrACBP : GLN 14: 0,08 nm, TYR 33: 0,05 nm, TYR 36: 0,05nm, LYS 37: 0,05 nm, LYS 59: 0,06 nm, TYR 78: 0,05 nm. Indicando menor flutuação devido a ligação da proteína com o ligante Isoxicam, sendo, portanto, favorável e indicando estabilidade. No final do gráfico apresentou pico de flutuação por ser considerada uma região terminal, em que se apresenta menos estável.

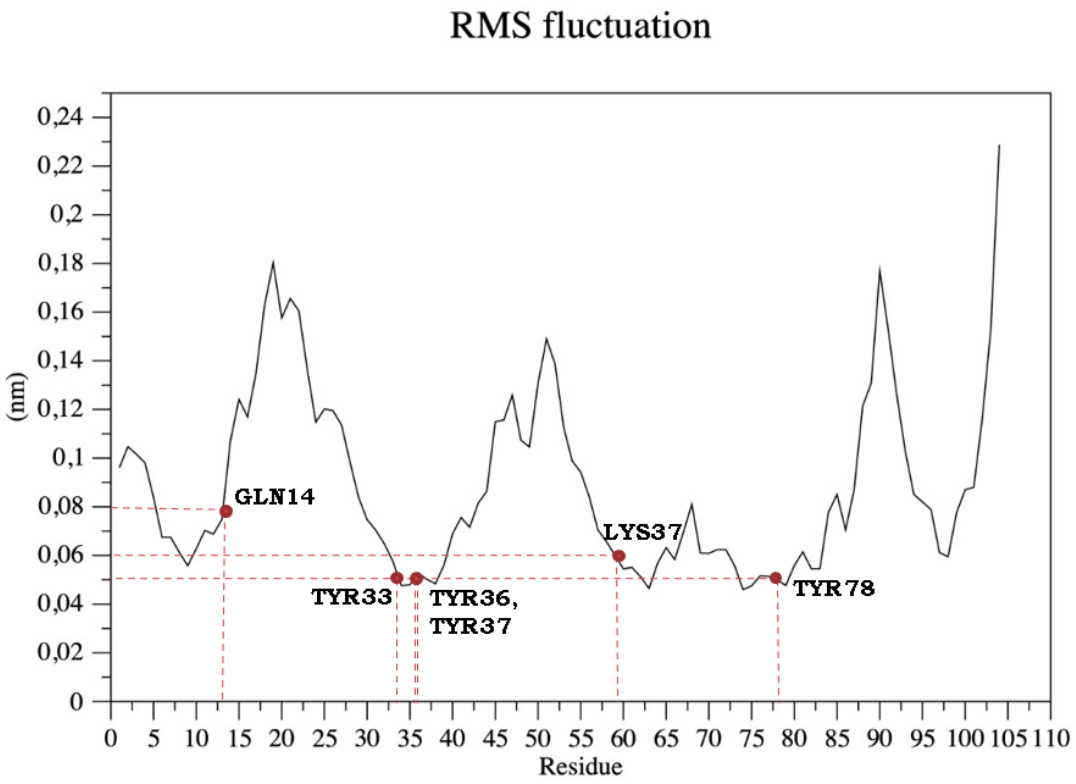

Figura 53: Gráfico da dinâmica molecular de estabilidade entre MrACBP e o ligante Isoxicam – Desvio da raiz quadrática média de flutuação. **Fonte:** Predição GROMACS 2020.

6. CONCLUSÃO

- Foi possível predizer a estrutura tridimensional da proteína de ligação acil-CoA (ACBP) do fungo basidiomiceto *Moniliophthora roreri* construído a partir da proteína ACBP do fungo *Moniliophthora perniciosa*, bem como, a estrutura 3D da ACBP do *Theobroma cacao* construído a partir da ACBP de *Orysa sativa.* Sendo estes identificados com melhor identidade nos servidores PDB e SWISS-MODEL Workspace, a partir da sequência de aminoácidos obtida de forma experimental.
- Na validação da qualidade tridimensional do modelo construído de MrACBP a partir do molde MpACBP, e da TcACBP a partir do molde OsACBP, foram encontrados os valores de identidade acima do que a literatura prediz para modelos aceitáveis. Os valores de QMQE e QMEAN foram consideravelmente acima do mínimo exigido. No gráfico de Ramachandran indicou que os resíduos das estruturas alvo localizam-se em zonas energeticamente mais favoráveis, considerando a modelagem válida. Desta forma foi possível predizer com confiabilidade o modelo estrutural da proteína ACBP de *M. roreri* e assim, melhor compreender a estrutura proteica.

- Os compostos químicos (ligantes) selecionados nesse estudo com os melhores valores de energia de afinidade na interação com MrACBP foram: o Ácido gambogico, Isoxicam, Meclociclina, Cloxacilina, Pregnenolona, porém não se mostraram seletivos na especificidade de interação, pois também tiveram afinidade de interação com TcACBP, sendo, portanto, indicados para mais análises.
- O Isoxicam se mostrou como um inibidor em potencial para MrACBP, mesmo ao interagir com menor energia de afinidade com TcACBP, mostrou maior energia de afinidade e seletividade em interagir com mais aminoácidos do sítio de MrACBP.
- A partir dos gráficos de Dinâmica Molecular de estabilidade podemos inferir que o RMSD da MrACBP com Isoxicam se mostrou instável, porém demonstrou estabilidade e afinidade nas médias de estimativa do potencial de energia, como também no gráfico RMSF dos resíduos do sítio da MrACBP e manteve um número considerável favorável de ligações de hidrogênios ao longo da trajetória de simulação. Portanto, sendo apto para prosseguir com testes *in vitro* e *in vivo* no intuito de atuar como novas perspectivas no controle da monília.

7. PERSPECTIVAS

- A pesquisa envolvendo proteínas da família ACBP em patógenos fúngicos é relativamente incomum, porém, explorar a estrutura e função fisiológica da proteína ACBP é de grande importância, uma vez que desempenham um papel importante durante o metabolismo de lipídios colaborando na função fisiológica normal da célula. Portanto, mais estudos para explorar a estrutura e função da ACBP em patógenos fúngicos e seu efeito na função fisiológica da célula podem contribuir para o desenvolvimento de novos métodos para reduzir o risco de patógenos.
- Avaliar a afinidade de interação do ligante com o TcACBP a partir da estrutura energética e a dinâmica molecular e verificar quanto a seletividade da interação dos aminoácidos com o sítio ativo. Além de propor novos trabalhos para compreender a funcionalidade da ACBP em *Theobroma cacao*, uma vez que neste trabalho foi identificado a presença de uma alça divergente do modelo construído a partir da OsACBP, sendo que essa região de alça pode ser considerado um local de ativação de efetores fitopatogênicos, de acordo com outras pesquisas que obtiveram resultados semelhantes, porém sua funcionalidade não está muito bem esclarecida.

8. REFERÊNCIAS BIBLIOGRÁFICAS

ABRAHAM, M. J.et al. GROMACS: High performance molecular simulations through multilevel parallelism from laptops to supercomputers**. Software X**, v. 1-2, p. 19-25, 2015.

ACRE, Mapas. Mapas do Brasil, **Guia Geográfico** – Estados brasileiros do Norte [s.l., s.n.] Disponível em <https://www.guiageo.com/acre.htm>. Acessado em: 6 de fevereiro de 2022.

ADASME et al. PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. **Nucleic Acids Research**, v. 49, n. W1, p. 1-5, 2021.

AIME, M. C.; PHILLIPS-MORA, W. The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. **Mycologia**, v. 97, n. 5, p. 1012–1022, 2005.

ALTSCHUL, S. F. et al. Ferramenta básica de pesquisa de alinhamento local. **J. Mol. Biol.**, v. 215, p. 403-410, 1990.

ALTSCHUL, S. F. et al. Gapped BLAST e PSI-BLAST: uma nova geração de programas de pesquisa de banco de dados de proteínas. **Nucleic Acids Research**, v. 25, n. 17, p. 3389-3402, 1997.

BAHIA, Governo do Estado. Agência Estadual de Defesa Agropecuária da Bahia – (ADAB). **ADAB desenvolve ações sincronizadas para manter a Bahia livre da monilíase do cacaueiro**. Salvador, Ba, publicado em 15 de julho de 2021a. Disponível em: <www.adab.ba.gov.br/2021/07/2639/ADAB-desenvolve-acoes-sincronizadas-para-manter-a-Bahia-livre-da-moniliase-do-cacaueiro.html>. Acessado em: 20 de ago. de 2021.

BAHIA, Governo do Estado. Secretaria de Agricultura, Pecuária, Irrigação, Pesca e Aquicultura - (SEAGRI). **NOTA TÉCNICA, A Bahia em alerta na prevenção à monilíase**. In. CARLOS, J.(gabinete do secretário), Salvador, Ba, publicado em 13 de julho de 2021b. Disponível em: < http://www.adab.ba.gov.br/arquivos/File/ASCOM_2021/DOCUMENTOS/15_07_2021_Nota_Teecnica_Moniliiase_versaao_final.pdf>. Acessado em 25 de ago. de 2021.

BAILEY, B.A. et al. Dynamic changes in pod and fungal physiology associated with the shift from biotrophy to necrotrophy during the infection of *Theobroma cacao* by *Moniliophthora roreri*. **Physiological and Molecular Plant Pathology**, v. 81, p. 84-96, 2013.

BAILEY, B. A. et al. *Moniliophthora roreri*, causal agent of cacao frosty pod rot. **Molecular Plant Pathology**, v. 19, n. 7, p. 1580–1594, 2018.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento – (MAPA); Secretaria de Defesa Agropecuária. Instrução Normativa nº 112, de 11 de dezembro de 2020. **In:** Insitui o Plano Nacional de Prevenção e Vigilância de *Moniliophthora roreri* – PNPV/Monilíase. **Diário Oficial da União**, publicado em 15 de dezembro de 2020, ed. 239, seção 1, p. 10. Disponível em: < https://www.in.gov.br/en/web/dou/-/instrucao-normativa-n-112-de-11-de-dezembro-de-2020-294054695>. Acessado em 20 de ago. de 2021.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento/Secretaria de Defesa Agropecuária. Portaria nº 372, de 3 de agosto de 2021a. **In:** Declara o estado do Acre como área sob quarentena para a praga quarentenária ausente *Moniliophthora roreri*. **Diário Oficial da União,** publicado em 05 de ago. de 2021a, ed. 147, seção 1, p. 3. Disponível em:

<https://in.gov.br/en/web/dou/-/portaria-n-372-de-3-de-agosto-de-2021-336315500> Acessado em: 25 de ago. de 2021.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento/Gabinete da Ministra. Portaria nº 249, de 4 de agosto de 2021b. **In:** Declara estado de emergência fitossanitária relativo ao risco iminente de introdução da praga quarentenária ausente *Moniliophthora roreri* nos estados do Acre, Amazonas e Rondônia. **Diário Oficial da União**, publicado em 06 de agosto de 2021b, ed. 148, seção 1, p.3. Disponível em: <https://www.in.gov.br/en/web/dou/-/portaria-mapa-n-249-de-4-de-agosto-de-2021-336635127> Acessado em: 25 de ago. de 2021.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento – (MAPA); Secretatia de Defesa Agropecuária; Departamento de Sanidade Vegetal e Insumos Agrícolas. Publicado em 05 de agosto de 2021c. **MANUAL DE PROCEDIMENTOS - Plano Nacional de Prevenção e Vigilãncia de *Moniliophthora roreri* PNPV/Monilíase**, versão 1.0, p. 35, 2021c. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/ceplac/publicacoes/moniliase/ManualProcedimentosMoniliaseVERSAO05agosto2021.pdf>. Acessado em: 25 de ago. de 2021.

BROOKS, B. R. et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. **Journal of Computational Chemistry**, v. 4, n. 2, p. 187–217, 1983.

CALIXTO, P. H. M. *et al*. Gene identification and comparative molecular modeling of a *Trypanosoma rangeli* major surface protease. **J Mol Model**, v. 19, n. 8, p. 3053-64, 2013.

CASE et al. AMBER 2020, University of California, San Francisco, 2020.

CIFERRI, R. & PARODI, E. Descrizione del fungo che causa la "Moniliasi" del cacao. **Phytopathologische Zeitschrift**, v. 6, p. 539-542, 1933.

CLARK, A. J. et al. Prediction of Protein–Ligand Binding Poses via a Combination of Induced Fit Docking and Metadynamics Simulations. Journal of Chemical Theory and Computation, v. 12, n. 6, p. 2990–2998, 2016.

COELHO, L. W. et al. Molecular mechanics application in inorganic chemistry. **Quimica Nova**, v. 22, n. 3, p. 396–404, 1999.

CORDERO-MORALES, J. F. et al. Molecular driving forces determining potassium channel slow inactivation**. Nature Structural & Molecular Biology**, v. 14, n.11, p. 1062–1069, 2007.

CORTES-CABRERA, A. et al. Ligand-Based Drug Discovery and Design. Em Cavasotto, C. N., editor, **In Silico Drug Discovery and Design: Theory, Methods, Challenges, and Applications**, cap. 4, p. 99-116. CRC Press, Londres, 2016.

COSTA, E.; GUIDOTTI, A. Diazepam binding inhibitor (DBI): A peptide with multiple biological actions. **Life Sciences**, vol. 49, n. 5, p. 325–344, 1991.

CRONSHAW, D. K. Fungicide application together with cultural practices to control cocoa diseases in Ecuador. **Tropical Agriculture** (Trinidad), v. 56, n. 2, p. 165–170, 1979.

DELEMOTTE, L. et al. Intermediate states of the Kv1.2 voltage sensor from atomistic molecular dynamics simulations. **Proceedings of the National Academy of Sciences**, v. 108, n. 15, p. 6109–6114, 2011.

DÍAZ-VALDERRAMA, J. R.; LEIVA-ESPINOZA, S. T.; CATHERINE-AIME, M. The history of cacao and its diseases in the Americas. **Phytopathology**, v. 110, n. 10, p. 1604–1619, 2020.

DI TOMMASO, P. et al. T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. **Nucleic Acids Research**, v. 39, n. SUPPL. 2, p. 13–17, 2011.

DOMINGUES, B. F. 3D-Pharma: uma ferramenta para triagem virtual baseada em fingerprints de farmacóforos. 2012. Tese (Doutorado em Bioinformática) Universidade Federal de Minas Gerais, Belo Horizonte, 193 p., 2012. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/BUBD-9DKHDA>. Acessado em: 02 out. 2021.

DU, X. et al. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. **International Journal of Molecular Sciences**, v. 17, n. 2, p. 1-34, 2016.

DURRANT, J. D.; McCAMMOM, J. A. Molecular dynamics simulations and drug discovery. **BMC Biology**, v. 9, n. 71, p 1-9, 2011.

EVANS, H. C.; STALPERS, J. A.; SAMSON, R. A.; BENNY, G. L On the taxonomy of Monilia roreri, an important pathogen of *Theobroma cacao* in South America. **Canadian Journal of Botany**, v. 56, p. 2528-2532, 1978.

EVANS, H. C. Pod rot of cacao caused by *Moniliophthora* (*Monilia*) *roreri*. **Phytopathological Papers**, v.24, p. 44, 1981.

EVANS, H. C. Invasive neotropical pathogens of tree crops. In: WATLING, R. et al. (eds). **Tropical Mycology**: **Vol.2. Micromycetes**. CABI Publishing, Wallingford, UK, cap. 7, p. 83-112, 2002.

EVANS, H. C.; HOLMES, K. A.; REID, A. P. Phylogeny of the frosty pod rot pathogen of cocoa. **Plant Pathology**, v. 52, n. 4, p. 476–485, 2003.

EVANS, H. Cacao diseases – The triology revisited. **Phytopathology** UK, v.97, p.1640-1643, 2007.

EVANS, H. C. Frosty Pod Rot (*Moniliophthora Roreri*). In: BAILEY, B.; MEINHARDT, L. (Eds.). Cacao Diseases - A History of Old Enemies and New Encounters. **Springer International Publishing Switzerland**, cap. 3, p.63-96, 2016.

FÆRGEMAN, N. J.; KNUDSEN, J. Role of long-chain fatty acyl-CoA esters in the regulation of metabolism and in cell signalling. **Biochemical Journal**, v. 323, n. 1, p. 1–12, 1997.

FRENKEL, Z. M.; TRIFONOV, E. N. Walking through the protein sequence space: towards new generation of the homology modeling. **Proteins: Structure, Function, and Bioinformatics**, v. 67, n. 2, p. 271-84, Mai., 2007.

FRITZLER, J. M.& ZHU, G. Novel anti-*Cryptosporidium* activity of known drugs identified by high-throughput screening against parasite fatty acyl-CoA binding protein (ACBP). **Journal of Antimicrobial Chemotherapy**, v. 67, n. 3, p. 609–617, 2011.

GASTEIGER, E. et al. The Proteomics Protocols Handbook. **The Proteomics Protocols Handbook**, p. 571–608, 2005.

GORDON, I. K.; KHANNA, C. Modeling opportunities in comparative oncology for drug development. **ILAR J**, v. 51, n. 3, p. 214- 20, dez., 2010.

GRAMACHO, K. P. et al. *Moniliophthora roreri* (Cif & Par.) (Agaricales: Marasmiaceae). In: FIDELIS, E. G. *et al*. (Eds). **Priorização de Pragas Quarentenárias ausentes no Brasil**. 1ª ed. Brasília, DF: Embrapa, cap. 17, p. 258-271, 2018.

GRIFFITH, G. W. et al. Witches' brooms and frosty pods: Two major pathogens of cacao. **New Zealand Journal of Botany**, v. 41, n. 3, p. 423–435, 2003.

GUO, Z. H. et al. The first plant acyl-CoA-binding protein structures: The close homologues OsACBP1 and OsACBP2 from rice. **Acta Crystallographica Section D: Structural Biology**, v. 73, n. 5, p. 438–448, 2017.

HUNTER C. A., & SANDERS, J. K. M. The nature of pi - pi interactions. **Journal of the American Chemical Society**, v. 112, n. 14, p. 5525–5534, 1990.

HANWELL, M. D. et al. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. Journal of Cheminformatics, v. 4, n. 1, p. 1-17, 2012.

HOLLINGSWORTH, S. A.; DROR, R. O. Molecular Dynamics Simulation for All. **Neuron**, v. 99, n. 6, p. 1129-1143, 2018.

HOSPITAL, A., et al. Molecular dynamics simulations: advances and applications. **Advances and Applications in Bioinformatics and Chemistry**, v. 2015, n. 8, p. 37-47, 2015.

HUANG, J., et al. CHARMM36m: na improved force field for folded and instrinsically disordered proteins. **Nature Methods**, v. 14, n. 1, p. 71-73, 2016.

HUEY, R., et al. A semiempirical free energy force field with charge-based desolvation. **Journal of Computational Chemistry**, v. 28, n. 6, p.1145-1152, 2007.

JO, S. et al. CHARMM-GUI: A web-based graphical user interface for CHARMM. **Journal of Computational Chemistry**, v. 29, n. 11, p. 1859–1865, 2008.

JO, S. et al. CHARMM-GUI 10 years for biomolecular modeling and simulation. **Journal of Computational Chemistry**, v. 38, n. 15, p. 1114–1124, 2016.

JOHNSON, E. S., et al. First report of *Moniliophthora roreri* causing frosty pod rot on Theobroma cacao in Jamaica. **New Disease Reports**, v. 36, p. 2, 2017.

KARPLUS, M.; McCAMMON, J. A. Molecular dynamics simulations of biomolecules. **Nature Structural Biology**, v. 9, n. 9, p. 646–652, 2002.

KO, J. et al. GalaxyWEB server for protein structure prediction and refinement. **Nucleic Acids Research**, v. 40, n. W1, p. 294–297, 2012.

KOPP, J.; SCHWEDE, T. The SWISSMODEL Repository of annotated threedimensional protein structure homology models. **Nucleic Acids Research**, v. 32, n. Database issue, p. D230-

D234, Jan., 2004.

KRAGELUND, B. B.; KNUDSEN, J.; POULSEN, F. M. Acyl-coenzyme A binding protein (ACBP). **Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids**, v. 1441, n. 2–3, p. 150–161, 1999.

LASKOWSI, R.A.et al. PROCHECK - a program to check the stereochemical quality of protein structures. **J. App. Cryst**, v. 26, p. 283-291,1993.

LAZIM, R.; SUH, D.; CHOI, S. Advances in Molecular Dynamics Simulations and Enhanced Sampling Methods for the Study of Protein Systems**. International Journal of Molecular Sciences**, v. 21, n. 17, p.1-20, 2020.

LEMKUL, J.A. "From Proteins to Perturbed Hamiltonians: A Suite of Tutorials for the GROMACS-2018 Molecular Simulation Package, v1.0" Virginia Tech Department of Biochemistry. ***Living J. Comp. Mol. Sci***, v. 1, n. 1, 2018. Disponível em: http://www.mdtutorials.com/gmx/complex/index.html . Acessado em: 30 de junho 2021.

LI, J., et al. Transient formation of water-conducting states in membrane transportes. **Proceedings of the National Academy of Sciences**, v. 110, n. 19, p. 7696-7701, 2013.

LINHARES, J. et al., Introdução à Triagem Virtual. **BIOINFO – Revista Brasileira de Bioinformática e Biologia Computacional.** Lagoa Santa: Editora Alfahelix, 1ª ed., v. 1, 2021. Disponível em: <**https://bioinfo.com.br/introducao-a-triagem-virtual/**>. Acessado em 20 de set. de 2021.

LIU, Y.; KE, M.; GONG, H. Protonation of Glu 135 Facilitates the Outward-to-Inward Structural Transition of Fucose Transporter. **Biophysical Journal**, v. 109, n. 3, p. 542–551, 2015.

MADHAVI SASTRY, G. et al. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. **Journal of Computer-Aided Molecular Design**, v. 27, n. 3, p. 221–234, 2013.

MANDRUP, S. et al. Acyl-CoA-binding protein/diazepam-binding inhibitor gene and pseudogenes. A typical housekeeping gene family. **Journal of Molecular Biology**, v. 228, n. 3, p. 1011–1022, 1992.

MANDRUP, S.; FÆRGEMAN, N. J.; KNUDSEN, J. Structure, Function, and Phylogeny of Acyl-CoA Binding Protein. **Cellular Proteins and Their Fatty Acids in Health and Disease**, cap. 8, p. 151–171, 2004.

MARCHLER-BAUER, A. et al.CDD: NCBI’s conserved domain database. Nucleic Acids Research, v. 43, n. D1, p. D222-D226, 2014.

MARELLI, J. P. et al. Chocolate under threat from old and new cacao diseases. **Phytopathology**, v. 109, n. 8, p. 1331–1343, 2019.

MARTÍ-RENOM, M. A., et al. Comparative Protein Structure Modeling of Genes and Genomes. **Annual Review of Biophysics and Biomolecular Structure**, v. 29, n. 1, p. 291–325, 2000.

MEINHARDT, L. W. et al. Genome and secretome analysis of the hemibiotrophic fungal pathogen, *Moniliophthora roreri*, which causes frosty pod rot disease of cacao: Mechanisms of the biotrophic and necrotrophic phases. **BMC Genomics**, v. 15, n. 1, 2014.

MITCHELL, A. L. et al. InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. **Nucleic Acids Research**, v. 47, n. D1, p. D351–D360, 2019.

MIRANDA, Morgana. Grande ameaça à produção do cacau na América Latina se aproxima do Brasil. In. Fernando (Ed.), Conselho Regional de Economia da Bahia – (CORECON, BA). **Jornal Correio**. Publicado: 25 de junho de 2018. Disponível em: <http://www.corecon-ba.org.br/grande-ameaca-a-producao-do-cacau-na-america-latina-se-aproxima-do-brasil/17839/>. Acessado em: 06/02/2022.

MOLCANOV, K, & KOJIC-PRODIC, B. Towards understanding π-stacking interactions between non-aromatic rings. **IUCrJ**, v. 6, n. 2, p. 156–166, 2019.

MONZANI, P. S. et al. A new topology of ACBP from Moniliophthora perniciosa. **Biochimica et Biophysica Acta - Proteins and Proteomics**, v. 1804, n. 1, p. 115–123, 2010.

MORRIS, G. M.; LIM-WILBY, M. Molecular Docking. In. Kukol, A. (eds) Methods in Molecular Biology, **Molecular Modeling of Proteins**, v. 443, cap. 19, p. 365-382, Humana Press, Totowa, NJ, 2008.

MORRIS, G. M., et al. Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. **Journal of Computational Chemistry**, v. 30, n. 16, p. 2785-2791, 2009.

NASCIMENTO, T. A. Valoração do potencial de impacto econômico com a introdução de *Moniliophthora roreri* no Brasil. Tese (Mestrado em Defesa Sanitária Vegetal) – Universidade Fedaral de Viçosa, MG, p.44, 2014. Disponível em: https://www.locus.ufv.br/bitstream/123456789/5875/1/texto%20completo.pdf Acessado em: 25 de ago. de 2021.

NELSON, D.L.; COX, M.M. Aminoácidos, Peptídeos e Proteínas. **Princípios de Bioquímica de Lehninger**. 6. ed. São Paulo: Artmed, Porto Alegre – RS, cap. 3, p. 75-114, 2014.

PETERSEN, T.N. et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, London, v. 8, p. 785-786, 2011.

PHILLIPS-MORA, W. **Origin, biogeography, genetic,diversity and taxonomic affinities of the cacao (*Theobroma cacao L.)* fungus *Monilophthora roreri* (Cif.) Evans et al. as determined using molecular, phytopathological and morpho-physiological evidence**. Tese (Doutorado) – Universidade de Reading, Deparment of Agricultural Botany, School of Plant Sciences, Reino Unido, p. 1-349, 2003.

PHILLIPS-MORA, W. et al. First report of Moniliophthora roreri causing frosty pod rot (= moniliasis disease) of cacao in Mexico. **Plant Pathology**, v. 55, p. 584, 2006a.

PHILLIPS-MORA, W. et al. First report of frosty pod rot (moniliasis disease) caused by *Moniliophthora roreri* on cacao in Belize. **Plant Pathology**, v. 55, p. 584, 2006b.

PHILLIPS-MORA, W.; AIME, M. C.; WILKINSON, M. J. Biodiversity and biogeography of the cacao (Theobroma cacao) pathogen *Moniliophthora roreri* in tropical America. **Plant**

Pathology, v. 56, n. 6, p. 911–922, 2007.

PHILLIPS-MORA, W. & WILKINSON, M.J. Frosty Pod of Cacao: A Disease with a Limited Geographic Range but Unlimited Potential for Damage. **Phytopathology**, v. 97:1644-1647, 2007.

PHILLIPS-MORA, W. et al. A. First report of frosty pod rot caused by *Moniliophthora roreri* on cacao in Bolivia**. New Disease Reports**, v. 31, p. 29, 2015.

PROVASI, D. et al. Ligand-Induced Modulation of the Free-Energy Landscape of G Protein-Coupled Receptors Explored by Adaptive Biasing Techniques. **PLoS Computational Biology**, v. 7, n. 10, p. 1-11, 2011.

RANGEL, F. S. Caracterização estrutural da ciclofilina do *Moniliophthora roreri* e o estudo de possíveis inibidores. Dissertação (mestrado em Genética e Biologia molecular) Universidade Estadual de Santa Cruz, Ilhéus, BA, 68 p., 2019.

RASMUSSEN, J. T.et al. Acyl-CoA-binding protein (ACBP) can mediate intermembrane acyl-CoA transport and donate acyl-CoA forβ-oxidation and glycerolipid synthesis. **Biochemical Journal**, 299(1), 165–170. doi:10.1042/bj2990165, 1994.

RORER, J. B. Diseases and pests of cacao in Ecuador and methods appropriate to the cultivation of cacao. Report, Agricultural Society of Ecuador, Guayaquil: **Imprenta do Diario Illustrado**, Ecuador, p 1-80, 1918.

ROSE, T. M.; SCHULTZ, E. R.; TODARO, G. J. Molecular cloning of the gene for the yeast homolog (ACB) of diazepam binding inhibitor/endozepine/acyl-CoA-binding protein. **Proceedings of the National Academy of Sciences**, vol. 89, n. 23, p. 11287–11291. Doi:10.1073/pnas.89.23.11287, 1992.

ROY, A.; KUCUKURAL, A.; ZHANG, Y. I-TASSER: A unified platform for automated protein structure and function prediction. **Nature Protocols**, v. 5, n. 4, p. 725–738, 2010.

SANCHEZ-MORA, F. D.; FIALLOS, F. R. G. *Moniliophthora roreri* (Cif y Par) Evans *et al*. en el cultivo de cacao. **Scientia Agropecuaria**, v. 3, p. 249–258, 2012.

SANTOS FILHO, O. A.; BICCA DE ALENCASTRO, R. Modelagem de proteínas por homologia. **Quimica Nova**, v. 26, n. 2, p. 253–259, 2003.

SCHRÖDINGER, L.L.C. The PyMOL Molecular Graphics System, Version 2.5.2, 2021.

SCHWEDE, T. et al. SWISS-MODEL: An automated protein homology-modeling server. **Nucleic Acids Research**, v. 31, n. 13, p. 3381–3385, 2003.

SGRILLO, R. Modelos Matemáticos para Simular a Introdução e Dispersão da Monilíase do Cacaueiro no Brasil. Relatório (Pós-Doutorado) - ESALQ/USP; Brasília, DF, p. 85 2010. CEPLAC/CEPLEC/MAPA, 2010. Disponível em: https://www.sgrillo.net/sysdyn/sgrillo_relatorio_final_monilia.pdf Acessado em: 25 de ago. de 2021.

SONG, Y. et al. High-resolution comparative modeling with RosettaCM. **Structure**, v. 21, n. 10, p. 1735–1742, 2013.

SUÁREZ, Y. J.; HERNÁNDEZ, F. A. Manejo de las enfermedades del cacao (*Theobroma cacao L.*) en Colombia, con énfasis en monilia (*Moniliophthora roreri*). Colombia. **Corpoica**. 90 p., 2010

TEILUM, K. et al. Different secondary structure elements as scaffolds for protein folding transition states of two homologous four-helix bundles. **Proteins: Structure, Function and Genetics**, v. 59, n. 1, p. 80–90, 2005.

TIRADO-GALLEGO, P. A.; LOPERA-ÁLVAREZ, A.; RÍOS-OSORIO, L. A. Estrategias de control de *Moniliophthora roreri* y *Moniliophthora perniciosa* en *Theobroma cacao L.*: revisión sistemática, **Corpoica Ciencia y Tecnología Agropecuaria**, v. 17, n. 3, p. 417–430, 2016.

TROTT, O.; OLSON, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, **Journal of Computational Chemistry**, v. 31, n. 2, p. 455-461, 2010.

VELOSO, W. N. P. Easyvs: uma ferramenta para triagem virtual mista baseada em alvo e ligante. Tese (Doutorado em Bioinformática) – Universidade Federal de Minas Gerais, Belo Horizonte, 71 p., 2019. Disponível em: http://hdl.handle.net/1843/30754. Acessado em: 21 jun. 2021.

VERLI, H. **Bioinformática da Biologia à flexibilidade molecular**, 1 ed. São Paulo: SBBq. 282 p., 2014.

WALTERS, W.; STAHL, M. T.; MURCKO, M. A. Virtual Screening - an overview. **Drug Discovery Today**, v. 3, n. 4, p. 160-178, 1998.

WARREN, G. L. et al. A critical assessment of docking programs and scoring functions. **Journal of Medicinal Chemistry**, v. 49, n. 20, p. 5912–5931, 2006.

WATERHOUSE, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic Acids Research,** v. 46, n. W1, p. W296-W303, 2018.

WU, S.; ZHANG, Y. LOMETS: A Local Meta-Threading-Server for Protein Structure Prediction. *Nucleic Acids Research, v. 35*, n. 10, p. 3375–3382, 2007.

ZHANG, Y. I-TASSER: Fully Automated Protein Structure Prediction in CASP8. *Proteins: Structure, Function, and Bioinformatics*, v. *77 (Suppl 9)*, p. 100–113, 2009.